

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2004, 13:45:21 ; Search time 739 Seconds

(without alignment)
10279.306 Million cell updates/sec

Title: US-10-063-735-127

Perfect score: 1505

Sequence: 1 cgggacgcgacccaagcag.....tcgcaaccccccaaaaaa 1505

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 3327077 seqs, 2523723180 residues

Word size : 10

Total number of hits satisfying chosen parameters: 3143743

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1505	100.0	1505	US-10-206-915-377	Sequence 377, App
2	1505	100.0	1505	US-10-199-670-377	Sequence 377, App
3	1505	100.0	1505	US-10-201-858-377	Sequence 377, App
4	1505	100.0	1505	US-10-205-890-377	Sequence 377, App
5	1505	100.0	1505	US-10-208-024-377	Sequence 377, App
6	1505	100.0	1505	US-10-201-853-377	Sequence 377, App
7	1505	100.0	1505	US-10-063-745-127	Sequence 127, App
8	1505	100.0	1505	US-10-063-513-127	Sequence 127, App
9	1505	100.0	1505	US-10-063-569-127	Sequence 127, App
10	1505	100.0	1505	US-10-063-551-127	Sequence 127, App
11	1505	100.0	1505	US-10-174-581-377	Sequence 377, App
12	1505	100.0	1505	US-10-176-483-377	Sequence 377, App
13	1505	100.0	1505	US-10-176-483-377	Sequence 377, App
14	1505	100.0	1505	US-10-176-749-377	Sequence 377, App

15	1505	100.0	1505	13	US-10-176-914-377	Sequence 377, App
16	1505	100.0	1505	13	US-10-176-915-377	Sequence 377, App
17	1505	100.0	1505	13	US-10-063-555-127	Sequence 127, App
18	1505	100.0	1505	13	US-10-063-563-127	Sequence 127, App
19	1505	100.0	1505	13	US-10-063-594-127	Sequence 127, App
20	1505	100.0	1505	13	US-10-063-553-127	Sequence 127, App
21	1505	100.0	1505	13	US-10-063-554-127	Sequence 127, App
22	1505	100.0	1505	13	US-10-176-484-377	Sequence 377, App
23	1505	100.0	1505	13	US-10-180-550-377	Sequence 377, App
24	1505	100.0	1505	13	US-10-183-014-377	Sequence 377, App
25	1505	100.0	1505	13	US-10-187-738-377	Sequence 377, App
26	1505	100.0	1505	13	US-10-187-740-377	Sequence 377, App
27	1505	100.0	1505	13	US-10-187-883-377	Sequence 377, App
28	1505	100.0	1505	13	US-10-194-363-377	Sequence 377, App
29	1505	100.0	1505	13	US-10-194-460-377	Sequence 377, App
30	1505	100.0	1505	13	US-10-194-463-377	Sequence 377, App
31	1505	100.0	1505	13	US-10-194-484-377	Sequence 377, App
32	1505	100.0	1505	13	US-10-195-884-377	Sequence 377, App
33	1505	100.0	1505	13	US-10-195-896-377	Sequence 377, App
34	1505	100.0	1505	13	US-10-196-744-377	Sequence 377, App
35	1505	100.0	1505	13	US-10-196-755-377	Sequence 377, App
36	1505	100.0	1505	13	US-10-196-757-377	Sequence 377, App
37	1505	100.0	1505	13	US-10-197-704-377	Sequence 377, App
38	1505	100.0	1505	13	US-10-197-710-377	Sequence 377, App
39	1505	100.0	1505	13	US-10-198-758-377	Sequence 377, App
40	1505	100.0	1505	13	US-10-198-766-377	Sequence 377, App
41	1505	100.0	1505	13	US-10-199-304-377	Sequence 377, App
42	1505	100.0	1505	13	US-10-199-309-377	Sequence 377, App
43	1505	100.0	1505	13	US-10-199-313-377	Sequence 377, App
44	1505	100.0	1505	13	US-10-199-456-377	Sequence 377, App
45	1505	100.0	1505	13	US-10-201-329-377	Sequence 377, App

ALIGNMENTS

RESULT 1
US-10-206-915-377
; Sequence 377, Application US/10206915
; Publication No. US20040029221A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430RIC513
; CURRENT APPLICATION NUMBER: US/10/206,915
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28

```

; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 377
; LENGTH: 1505
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-206-915-377

Query Match      100.0%; Score 1505; DB 13; Length 1505;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGGATCGGACCAAGCAGGTCCGGCGCGCGCGCAGGAGCGCGGGCGTCAGCTCC 60
DB 1 CGCGGATCGGACCAAGCAGGTCCGGCGCGCGCGCAGGAGCGCGGGCGTCAGCTCC 60

QY 61 TCGACCCCGGTGTCGGGCTAGTCCAGCGAGCGCGAGCGGGCCCATGGCCAGG 120
DB 61 TCGACCCCGGTGTCGGGCTAGTCCAGCGAGCGCGAGCGGGCCCATGGCCAGG 120

QY 121 CCCGGCATGAGCGGTGCGCGCACCGGCTGCGCGTGGTGACGGGGGGCCATC 180
DB 121 CCCGGCATGAGCGGTGCGCGCACCGGCTGCGCGTGGTGACGGGGGGCCATC 180

QY 181 GCGCGGCGGTGGCGCGCGCCCTGGTCCAGCAGGAGCTGAAGGTGGTGGGCTGGCCCGC 240
DB 181 GCGCGGCGGTGGCGCGCGCCCTGGTCCAGCAGGAGCTGAAGGTGGTGGGCTGGCCCGC 240

QY 241 ACTGTGGCAACATCGAGGAGCTGGCTGCTGAATGTAAAGATGACGAGTACCCCGGACT 300
DB 241 ACTGTGGCAACATCGAGGAGCTGGCTGCTGAATGTAAAGATGACGAGTACCCCGGACT 300

QY 301 TTGATCCCTACAGATGACCTATCAATGAAGAGGACATCTTCCATGTTCTCAGCT 360
DB 301 TTGATCCCTACAGATGACCTATCAATGAAGAGGACATCTTCCATGTTCTCAGCT 360

QY 361 ATCCGTTCTCAGCAGCGGTGTAGACATCTGCATCAACATGCTGGCTTGGCCCGGCT 420
DB 361 ATCCGTTCTCAGCAGCGGTGTAGACATCTGCATCAACATGCTGGCTTGGCCCGGCT 420

QY 421 GACACCCCTGCTCTCAGGACGACACCATGCTGCAATGCTGCAATGCTGCAATGCTGCAAT 480
DB 421 GACACCCCTGCTCTCAGGACGACACCATGCTGCAATGCTGCAATGCTGCAATGCTGCAAT 480

QY 481 GCGCTCAGCATCTGACACGCGGAGCGCTACAGTCCATGAAGGAGCGGAATGTGACGAT 540
DB 481 GCGCTCAGCATCTGACACGCGGAGCGCTACAGTCCATGAAGGAGCGGAATGTGACGAT 540

QY 541 GGGCACATCTTAACATCAATAGCATGCTGGCCACCGAGTGTACCCCTGCTGTGACC 600
DB 541 GGGCACATCTTAACATCAATAGCATGCTGGCCACCGAGTGTACCCCTGCTGTGACC 600

QY 601 CACTTCTATAGTCCACCAAGATGTCGCTGCTGACAGGAGTACGAGGAGTACGAGGAG 660
DB 601 CACTTCTATAGTCCACCAAGATGTCGCTGCTGACAGGAGTACGAGGAGTACGAGGAG 660

QY 661 CTTCCGGAGGCGCCAGACCCACATCCGAGCGAGTGCATCTTCCAGGTGTGTGGAGACA 720
DB 661 CTTCCGGAGGCGCCAGACCCACATCCGAGCGAGTGCATCTTCCAGGTGTGTGGAGACA 720

QY 721 CAATTGCGCTTCAAACTCCACGACGAGGACCTGAGAGGAGTGCACCTATGAGCAA 780
DB 721 CAATTGCGCTTCAAACTCCACGACGAGGACCTGAGAGGAGTGCACCTATGAGCAA 780

QY 781 ATGAAGTGTCTAAACCCGAGAGTGTGGCCGAGGCTGTATCTAGCTCTCAGACCCCC 840
DB 781 ATGAAGTGTCTAAACCCGAGAGTGTGGCCGAGGCTGTATCTAGCTCTCAGACCCCC 840

QY 841 GCACACATCCAGATTGGAGACATCCAGATGAGGCCCGCAGGACGAGTGTGACCTGACTG 900

```

```

DB 841 GCACACATCCAGATTGGAGACATCCAGATGAGGCCCGCAGGAGGAGTGCACCTAGTACTG 900
QY 901 TGGGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 960
DB 901 TGGGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 960
QY 961 TGTGATTCTTGATCAGGGATACCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1020
DB 961 TGTGATTCTTGATCAGGGATACCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1020
QY 1021 ATTGTTTGGATTTTATATCATCTTGTCAAAATGCTTCAAGTTGTAATGTGAAAAATG 1080
DB 1021 ATTGTTTGGATTTTATATCATCTTGTCAAAATGCTTCAAGTTGTAATGTGAAAAATG 1080
QY 1081 GGCCTGGGAAAGGAGGTGGTGTCCCTAAATGTTTAACTTGTAACTTGTCTTCTGCCCC 1140
DB 1081 GGCCTGGGAAAGGAGGTGGTGTCCCTAAATGTTTAACTTGTAACTTGTCTTCTGCCCC 1140
QY 1141 TGGGCACTTGGCCTTGTCTGCTCTCAGTGTCTTCCCTTTGACATGGGAAAGGAGTTGTG 1200
DB 1141 TGGGCACTTGGCCTTGTCTGCTCTCAGTGTCTTCCCTTTGACATGGGAAAGGAGTTGTG 1200
QY 1201 GCCAAAATCCCATCTTCTTGGACCTCAAAGTCTGTGGCTCAGGGCTGGGTGGCAGAGG 1260
DB 1201 GCCAAAATCCCATCTTCTTGGACCTCAAAGTCTGTGGCTCAGGGCTGGGTGGCAGAGG 1260
QY 1261 GAGCCTTCACTTATATCTGTGTGTATCCAGGGCTCCAGACTTCTCTCTCTCTCTCTCT 1320
DB 1261 GAGCCTTCACTTATATCTGTGTGTATCCAGGGCTCCAGACTTCTCTCTCTCTCTCTCT 1320
QY 1321 CCAGTGGACCTCTCCCTTATCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1380
DB 1321 CCAGTGGACCTCTCCCTTATCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1380
QY 1381 TCTTGTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1440
DB 1381 TCTTGTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1440
QY 1441 GGGCCTGGCCAGTGGATTTTCATGGTATCAATTAATAAGAAAAATGCAACCAAAAAA 1500
DB 1441 GGGCCTGGCCAGTGGATTTTCATGGTATCAATTAATAAGAAAAATGCAACCAAAAAA 1500
QY 1501 AAAAA 1505
DB 1501 AAAAA 1505

```

```

RESULT 2
US-10-199-670-377
; Sequence 377, Application US/10199670
; Publication No. US20040033560A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C401
; CURRENT APPLICATION NUMBER: US/10/199,670
; CURRENT FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18

```



```

; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 377
; LENGTH: 1505
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-199-670-377

Query Match      100.0%; Score 1505; DB 13; Length 1505;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGGATCGAACCCAGAGAGGTGGCGCGCGCGCGAGAGAGCGCGCGCGTCC 60
Db 1 CGCGGATCGAACCCAGAGAGGTGGCGCGCGCGCGAGAGAGCGCGCGCGTCC 60

Qy 61 TCGACCCCGCTGCGGGCTAGTCAGAGCGAGCGAGCGCGCGCGTCCAGTCC 120
Db 61 TCGACCCCGCTGCGGGCTAGTCAGAGCGAGCGAGCGCGCGCGTCCAGTCC 120

Qy 121 CCGCGCATGAGCGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
Db 121 CCGCGCATGAGCGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180

Qy 181 GCGCGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
Db 181 GCGCGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240

Qy 241 ACTGTGGGCAACATCGAGAGCGTGGCTGCTGAATGAAGAGTGAAGAGT 300
Db 241 ACTGTGGGCAACATCGAGAGCGTGGCTGCTGAATGAAGAGTGAAGAGT 300

Qy 301 TTGATCCCTACAGATGTGACCTATCAATGAAGAGGACATCTCTCCATGTTCT 360
Db 301 TTGATCCCTACAGATGTGACCTATCAATGAAGAGGACATCTCTCCATGTTCT 360

Qy 361 ATCCGTTCTCAGCACAGCGGTAGACATCTGCATCAACAATGCTGGCTTGG 420
Db 361 ATCCGTTCTCAGCACAGCGGTAGACATCTGCATCAACAATGCTGGCTTGG 420

Qy 421 GACACCTGCTCTCAGCACAGCGGTGGTGGAGGACATGTTCAATGTAAGT 480
Db 421 GACACCTGCTCTCAGCACAGCGGTGGTGGAGGACATGTTCAATGTAAGT 480

Qy 481 GCCTCTCAGCATCTGCACACCGGAGCGTACAGTCCATGAAGAGCGGAATG 540
Db 481 GCCTCTCAGCATCTGCACACCGGAGCGTACAGTCCATGAAGAGCGGAATG 540

Qy 541 GGGCACATCAATTAACATAGCATGTCTGGCCACCGAGTGTACCCCTGTCTG 600
Db 541 GGGCACATCAATTAACATAGCATGTCTGGCCACCGAGTGTACCCCTGTCTG 600

Qy 601 CACTTCTATAGTGCACCAATAGTACCGTCACTGCGCTGACAGAGGAGTGA 660
Db 601 CACTTCTATAGTGCACCAATAGTACCGTCACTGCGCTGACAGAGGAGTGA 660

Qy 661 CTTCCGGAGGCCAGACCCACATCCGAGCGACAGTGCATCTCTCCAGGTG 720

```

RESULT 3

```

US-10-201-858-377
; Sequence 377, Application US/10201858
; Publication No. US2004003837A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria

```

```

Db 661 CTTCCGGAGGCCAGACCCACATCCGAGCGACGTCGATCTCTCAGGTGTTGGAGACA 720
Qy 721 CAATTTCGCTTCAAACTCCAAGAGGACCTCTGAGAAAGGAGCTGCCACCTATGAGCAA 780
Db 721 CAATTTCGCTTCAAACTCCAAGAGGACCTCTGAGAAAGGAGCTGCCACCTATGAGCAA 780
Qy 781 ATGAAGTGTCTCAAAACCCGAGGATGTGGCGAGGCTGTTATCTACGTCTCTCAGCACCC 840
Db 781 ATGAAGTGTCTCAAAACCCGAGGATGTGGCGAGGCTGTTATCTACGTCTCTCAGCACCC 840
Qy 841 GCACACATCCAGATTTGGAGACATCCAGATGAGGCCACCGAGAGGAGTGAAGTGAAGT 900
Db 841 GCACACATCCAGATTTGGAGACATCCAGATGAGGCCACCGAGAGGAGTGAAGTGAAGT 900
Qy 901 TGGAGGCTCTCTCTCCCTCCCAACCTTCATGCGCTTGCCTCTCTCTCTCTCTCTCTCTCT 960
Db 901 TGGAGGCTCTCTCTCCCTCCCAACCTTCATGCGCTTGCCTCTCTCTCTCTCTCTCTCTCT 960
Qy 961 TGTGTATTTCTGGATCAGCGGATACCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1020
Db 961 TGTGTATTTCTGGATCAGCGGATACCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1020
Qy 1021 ATTGTGTTGAGATTTTATATCATCTTGTCAAAATTTGCTTCAGTTGTAAATGTGAAATG 1080
Db 1021 ATTGTGTTGAGATTTTATATCATCTTGTCAAAATTTGCTTCAGTTGTAAATGTGAAATG 1080
Qy 1081 GGCTGGGAAAGAGGAGTGTCTCTAAATTTTACTTCTTAACTTCTTCTCTCTCTCTCTCTCT 1140
Db 1081 GGCTGGGAAAGAGGAGTGTCTCTAAATTTTACTTCTTAACTTCTTCTCTCTCTCTCTCTCT 1140
Qy 1141 TGGGCACTTGGCTTGTCTGCTCTCAGTGTCTTCCCTTTGACATGCGGAAAGAGTGTGTG 1200
Db 1141 TGGGCACTTGGCTTGTCTGCTCTCAGTGTCTTCCCTTTGACATGCGGAAAGAGTGTGTG 1200
Qy 1201 GCCAAATCCCATCTTCTTGACCTCAACGCTCTGTGGCTCAGGCTGCGGCTGGCAGAGG 1260
Db 1201 GCCAAATCCCATCTTCTTGACCTCAACGCTCTGTGGCTCAGGCTGCGGCTGGCAGAGG 1260
Qy 1261 GAGGCTTCACTTATATCTGTGTATFCAAGGCTCCAGACTTCTCTCTCTCTCTCTCTCTCT 1320
Db 1261 GAGGCTTCACTTATATCTGTGTATFCAAGGCTCCAGACTTCTCTCTCTCTCTCTCTCTCT 1320
Qy 1321 CCACCTGCACCTCTCCCTCTTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1380
Db 1321 CCACCTGCACCTCTCCCTCTTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1380
Qy 1381 TCTTGTCCCTCTCTGGGTCATCCCTCCACTCTGACTCTGACTATGCGAGCAACACCA 1440
Db 1381 TCTTGTCCCTCTCTGGGTCATCCCTCCACTCTGACTCTGACTATGCGAGCAACACCA 1440
Qy 1441 GGGCTGGCCAGTGGATTTGATGTCATTTAAAGAAAGAAATCGCAACCAAAAAA 1500
Db 1441 GGGCTGGCCAGTGGATTTGATGTCATTTAAAGAAAGAAATCGCAACCAAAAAA 1500
Qy 1501 AAAAA 1505
Db 1501 AAAAA 1505

```


Db 1201 GCCAAATCCCATCTTCTTGACCTCAACGCTGTGGCTCAGGGCTGGGGTGGCAGAGG 1260
 Qy 1261 GAGCCTTACCTTATATCTGTGTGTATATCCAGGGCTCAGAGCTTCTCTCTGGCTGC 1320
 Db 1261 GAGCCTTACCTTATATCTGTGTGTATATCCAGGGCTCAGAGCTTCTCTCTGGCTGC 1320
 Qy 1321 CCCACTGCACCTCTCCGCCCTTATCTATCTCTTCTGGCTCCCGAGCCAGTCTTGGCT 1380
 Db 1321 CCCACTGCACCTCTCCGCCCTTATCTATCTCTTCTGGCTCCCGAGCCAGTCTTGGCT 1380
 Qy 1381 TCTTGTCCCTCTCCGGGTATCCCTCCACTCTGACTCTGACTATGGCAGCAGAACACCA 1440
 Db 1381 TCTTGTCCCTCTCCGGGTATCCCTCCACTCTGACTCTGACTATGGCAGCAGAACACCA 1440
 Qy 1441 GGGCTGGCCAGTGGATTTTATGGTGATCAATTTAAAGAAAAAATGCAACCAAAAAA 1500
 Db 1441 GGGCTGGCCAGTGGATTTTATGGTGATCAATTTAAAGAAAAAATGCAACCAAAAAA 1500
 Qy 1501 AAAAA 1505
 Db 1501 AAAAA 1505

RESULT 6

US-10-201-853-377
 ; Sequence 377, Application US/10201853
 ; Publication No. US20040053358A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; TITLE OF INVENTION: ACIDS ENCODING THE SAME
 ; FILE REFERENCE: P3430R1C465
 ; CURRENT APPLICATION NUMBER: US/10/201,853
 ; CURRENT FILING DATE: 2002-07-23
 ; PRIOR APPLICATION NUMBER: 10/052586
 ; PRIOR FILING DATE: 2002-01-15
 ; PRIOR APPLICATION NUMBER: 60/059263
 ; PRIOR FILING DATE: 1997-09-18
 ; PRIOR APPLICATION NUMBER: 60/059266
 ; PRIOR FILING DATE: 1997-09-18
 ; PRIOR APPLICATION NUMBER: 60/062250
 ; PRIOR FILING DATE: 1997-10-17
 ; PRIOR APPLICATION NUMBER: 60/063120
 ; PRIOR FILING DATE: 1997-10-24
 ; PRIOR APPLICATION NUMBER: 60/063121
 ; PRIOR FILING DATE: 1997-10-24
 ; PRIOR APPLICATION NUMBER: 60/063486
 ; PRIOR FILING DATE: 1997-10-21
 ; PRIOR APPLICATION NUMBER: 60/063540
 ; PRIOR FILING DATE: 1997-10-28
 ; PRIOR APPLICATION NUMBER: 60/063541
 ; PRIOR FILING DATE: 1997-10-28
 ; PRIOR APPLICATION NUMBER: 60/063544
 ; PRIOR FILING DATE: 1997-10-28
 ; Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 612
 ; SEQ ID NO 377
 ; LENGTH: 1505
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 US-10-201-853-377

Query Match 100.0%; Score 1505; DB 13; Length 1505;

Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 CGCGATCGAACCCAAACAGAGTGGCGCGCGCGCAGAGAGCGCGCGCGTCTCAGTCC 60
 Db 1 CGCGATCGAACCCAAACAGAGTGGCGCGCGCGCAGAGAGCGCGCGCGTCTCAGTCC 60
 Qy 61 TCGACCCCTGTGTGGGCTAGTCCAGCGAGCGGCGCGCGTGGGCGGCGGCGGCGGCGG 120
 Db 61 TCGACCCCTGTGTGGGCTAGTCCAGCGAGCGGCGCGCGTGGGCGGCGGCGGCGGCGG 120
 Qy 121 CGCGCATGAGCGGTGGCGCGCGCGTGGTGAAGCGGCGCGTGGGCGGCGGCGGCGG 180
 Db 121 CGCGCATGAGCGGTGGCGCGCGCGTGGTGAAGCGGCGCGTGGGCGGCGGCGGCGG 180
 Qy 181 GGCAGCGCGTGGCGCGCGCGTGGTCCAGCAGGAGCTGAAGTGGTGGGCGGCGGCGG 240
 Db 181 GGCAGCGCGTGGCGCGCGCGTGGTCCAGCAGGAGCTGAAGTGGTGGGCGGCGGCGG 240
 Qy 241 ACTGTGGGCAACATCGAGGAGCTGGCTGCTGAATGTAAGTGAAGTGAAGTGAAGTGA 300
 Db 241 ACTGTGGGCAACATCGAGGAGCTGGCTGCTGAATGTAAGTGAAGTGAAGTGAAGTGA 300
 Qy 301 TTGATCCCTACAGATGTGACCTATCAATGAAGAGCAGATCCCTCTCATGTTCTCAGCT 360
 Db 301 TTGATCCCTACAGATGTGACCTATCAATGAAGAGCAGATCCCTCTCATGTTCTCAGCT 360
 Qy 361 ATCCGTTCTCAGCACAGCGGTGTAGACATCTGCATCAACATGCTGGCTTGGCGCGGCT 420
 Db 361 ATCCGTTCTCAGCACAGCGGTGTAGACATCTGCATCAACATGCTGGCTTGGCGCGGCT 420
 Qy 421 GACACCTCTCTCAGCAGCAGCAGTGGTGGAGGAGCAGATGTTCAATGTGAAGTGGTG 480
 Db 421 GACACCTCTCTCAGCAGCAGCAGTGGTGGAGGAGCAGATGTTCAATGTGAAGTGGTG 480
 Qy 481 GCCCTCAGCATCTGCACACGGGAAGCCTACAGTCCATGAAGGAGCGGAATGTGGACGAT 540
 Db 481 GCCCTCAGCATCTGCACACGGGAAGCCTACAGTCCATGAAGGAGCGGAATGTGGACGAT 540
 Qy 541 GGGCACATCAATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGTCTGTGACC 600
 Db 541 GGGCACATCAATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGTCTGTGACC 600
 Qy 601 CACTTCTATAGTGCACCAAGTATGCGCTCACTCCGAGCCAGTGTACCCCTGTCTGTGACC 660
 Db 601 CACTTCTATAGTGCACCAAGTATGCGCTCACTCCGAGCCAGTGTACCCCTGTCTGTGACC 660
 Qy 661 CTTGGGAGGCCAGAGCCACATCCGAGCCAGTGTGAGGAGCAGTGTGAGGAGCAGAG 720
 Db 661 CTTGGGAGGCCAGAGCCACATCCGAGCCAGTGTGAGGAGCAGTGTGAGGAGCAGAG 720
 Qy 721 CAATTGCGCTTCAAACTCCACGACCAAGGACCTTGAAGGAGCAGTGTGAGGAGCAG 780
 Db 721 CAATTGCGCTTCAAACTCCACGACCAAGGACCTTGAAGGAGCAGTGTGAGGAGCAG 780
 Qy 781 ATGAAGTGTCTCAAACTCCACGAGGAGTGTGGCGAGGCTGTATCTACGCTCTCAGCAC 840
 Db 781 ATGAAGTGTCTCAAACTCCACGAGGAGTGTGGCGAGGCTGTATCTACGCTCTCAGCAC 840
 Qy 841 GCACATCCAGATTTGAGAGCATCCAGATGAGGCGCCACGAGGAGTGAACCTAGTGA 900
 Db 841 GCACATCCAGATTTGAGAGCATCCAGATGAGGCGCCACGAGGAGTGAACCTAGTGA 900
 Qy 901 TGGGAGCT 960
 Db 901 TGGGAGCT 960
 Qy 961 TGTGATTTCTGATTCACGGGATACCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1020
 Db 961 TGTGATTTCTGATTCACGGGATACCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1020
 Qy 1021 ATTTGTTTGGATTTTATATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1080

Db 1021 ATTTGTTTGAGATTTTATATCATCTTGTCAAAATGCTTCAAGTTGTAATGTGAAAATG 1080
Qy 1081 GGCTGGGAAAGAGGTGGTCCCTAATGTTTACTTGTAACTTGTCTTGTGCCCC 1140
Db 1081 GGCTGGGAAAGAGGTGGTCCCTAATGTTTACTTGTAACTTGTCTTGTGCCCC 1140
Qy 1141 TGGGCACATTGGCTTTGTCTGTCTCAGTGTCTTCCCTTTGACATGGGAAAGGAGTTGTG 1200
Db 1141 TGGGCACATTGGCTTTGTCTGTCTCAGTGTCTTCCCTTTGACATGGGAAAGGAGTTGTG 1200
Qy 1201 GCCAAATCCCAATCTTCTTGCACCTCAACGTCTGTGGCTCAGGGCTGGGGTGGCAGAGG 1260
Db 1201 GCCAAATCCCAATCTTCTTGCACCTCAACGTCTGTGGCTCAGGGCTGGGGTGGCAGAGG 1260
Qy 1261 GAGGCTTCACTTATATCTGTGTGTTATCCAGGGCTCCAGACTTCCCTCTGCTGCTGC 1320
Db 1261 GAGGCTTCACTTATATCTGTGTGTTATCCAGGGCTCCAGACTTCCCTCTGCTGCTGC 1320
Qy 1321 CCCACTGACCTCTCCCTCTTATCTATCTCTCTCGGCTCCCGAGCCAGTCTTGGCT 1380
Db 1321 CCCACTGACCTCTCCCTCTTATCTATCTCTCTCGGCTCCCGAGCCAGTCTTGGCT 1380
Qy 1381 TCTTGTCCCTCTCGGGTCACTCCCTCCACTCTGACTCTGACTATGGCAGCAGAACCA 1440
Db 1381 TCTTGTCCCTCTCGGGTCACTCCCTCCACTCTGACTCTGACTATGGCAGCAGAACCA 1440
Qy 1441 GGGCTGGCCAGTGGATTTATGTTGATCAATTAAGAAAGAAATCCGACCCAAAGAAA 1500
Db 1441 GGGCTGGCCAGTGGATTTATGTTGATCAATTAAGAAAGAAATCCGACCCAAAGAAA 1500
Qy 1501 AAAAA 1505
Db 1501 AAAAA 1505

RESULT 7

US-10-063-745-127
; Sequence 127, Application US/10063745
; Publication No. US20040058411A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,745
; Prior Filing DATE: 2002-05-09
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 127
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-063-745-127

Query Match 100.0%; Score 1505; DB 13; Length 1505;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CGCGGATCGGACCAAGCAGGTCCGGCGCGCGGAGAGCGCGCGGGGCTCAGCTCC 60
Db 1 CGCGGATCGGACCAAGCAGGTCCGGCGCGCGGAGAGCGCGCGGGGCTCAGCTCC 60
Qy 61 TCGACCCCGGTGTCGGGTAGTCCAGAGCGGAGCGGCGGCTGCGCCCATGGCCAGG 120
Db 61 TCGACCCCGGTGTCGGGTAGTCCAGAGCGGAGCGGCGGCTGCGCCCATGGCCAGG 120

Qy 121 CC CGCATGGA GCGGTGGCGGACCGGCTGGCGCTGGTGACGGGGGCCCTCGGGGGGCATC 180
Db 121 CC CGCATGGA GCGGTGGCGGACCGGCTGGCGCTGGTGACGGGGGCCCTCGGGGGGCATC 180
Qy 181 GGCGCGGCGGTGGCGCGGGCCCTGGTCCAGCAGGAGACTGAAGGTGGTGGGCTGCGCCCGC 240
Db 181 GGCGCGGCGGTGGCGCGGGCCCTGGTCCAGCAGGAGACTGAAGGTGGTGGGCTGCGCCCGC 240
Qy 241 ACTGTGGGCAACATCGAGGAGCTGGCTGCTGAATGTAAAGTGCAGGCTACCCGGGACT 300
Db 241 ACTGTGGGCAACATCGAGGAGCTGGCTGCTGAATGTAAAGTGCAGGCTACCCGGGACT 300
Qy 301 TTGATCCCTTACAGATGTGACCTATCAAAATGAAGAGGACATCCTCTCCATGTTCTCAGCT 360
Db 301 TTGATCCCTTACAGATGTGACCTATCAAAATGAAGAGGACATCCTCTCCATGTTCTCAGCT 360
Qy 361 ATCCGTTCTCAGCACAGCGGTGTAGACATCTGCAATCAATGTGGCTTGGCCCGGCT 420
Db 361 ATCCGTTCTCAGCACAGCGGTGTAGACATCTGCAATCAATGTGGCTTGGCCCGGCT 420
Qy 421 GACACCTGTCTCAGGAGCACCAGTGGTTGGAAGGACATGTTCAATGTGAACGTGCTG 480
Db 421 GACACCTGTCTCAGGAGCACCAGTGGTTGGAAGGACATGTTCAATGTGAACGTGCTG 480
Qy 481 GCGCTCAGCATCTGCACAGGAAAGCCCTACAGTCCATGAAGAGCGGAATGTGGACAT 540
Db 481 GCGCTCAGCATCTGCACAGGAAAGCCCTACAGTCCATGAAGAGCGGAATGTGGACAT 540
Qy 541 GGGCACAATCAATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTCTGTGACC 600
Db 541 GGGCACAATCAATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTCTGTGACC 600
Qy 601 CACTTCTATAGTCCACCAAGTATGCGCTCACTGGCTGACAGAGGAGTCAAGCAAGAG 660
Db 601 CACTTCTATAGTCCACCAAGTATGCGCTCACTGGCTGACAGAGGAGTCAAGCAAGAG 660
Qy 661 CTTGGGAGGCGGAGACCCACATCCGAGCCACGTCGATCTCTCCAGTGTGGTGAGACA 720
Db 661 CTTGGGAGGCGGAGACCCACATCCGAGCCACGTCGATCTCTCCAGTGTGGTGAGACA 720
Qy 721 CAATTCGCTTCAAACTCCACGAGGACCCCTGAGAAGGAGCTGCCACCTATGAGCAA 780
Db 721 CAATTCGCTTCAAACTCCACGAGGACCCCTGAGAAGGAGCTGCCACCTATGAGCAA 780
Qy 781 ATGAAGTGTCTCAAAACCGGAGGATGTGGCCGAGGCTGTATCTACGCTCTCAGCACCC 840
Db 781 ATGAAGTGTCTCAAAACCGGAGGATGTGGCCGAGGCTGTATCTACGCTCTCAGCACCC 840
Qy 841 GCACATCCAGATTGGAGACATCCAGTGGGCGGCGGAGGAGTGCCTAGTGTGACTG 900
Db 841 GCACATCCAGATTGGAGACATCCAGTGGGCGGCGGAGGAGTGCCTAGTGTGACTG 900
Qy 901 TGGGAGCTCTCTCCCTCCCGACCTTCAATGCTTGCCTCTCGCTCTGATTTAGG 960
Db 901 TGGGAGCTCTCTCCCTCCCGACCTTCAATGCTTGCCTCTCGCTCTGATTTAGG 960
Qy 961 TGTGATTCTTGGATCAGGAGTACCACTTCTCTCCACACCCCGACAGGGGCTAGAAA 1020
Db 961 TGTGATTCTTGGATCAGGAGTACCACTTCTCTCCACACCCCGACAGGGGCTAGAAA 1020
Qy 1021 ATTTGTTTGAGATTTTATATCATCTTGTCAAAATGCTTCAAGTTGTAATGTGAAAATG 1080
Db 1021 ATTTGTTTGAGATTTTATATCATCTTGTCAAAATGCTTCAAGTTGTAATGTGAAAATG 1080
Qy 1081 GGCTGGGAAAGGAGGTGGTGTCCCTAATGTTTACTTGTAACTTGTCTTGTGCCCC 1140
Db 1081 GGCTGGGAAAGGAGGTGGTGTCCCTAATGTTTACTTGTAACTTGTCTTGTGCCCC 1140
Qy 1141 TGGGCACATTGGCTTTGTCTGCTCTCAGTGTCTTCCCTTTGACATGGGAAAGGAGTTGTG 1200
Db 1141 TGGGCACATTGGCTTTGTCTGCTCTCAGTGTCTTCCCTTTGACATGGGAAAGGAGTTGTG 1200


```

QY 1201 GCCAAATCCCATCTTCTGCACTCAACGCTGTGCTCAGGCTGGGTGGCAGG 1260
Db |||||
QY 1201 GCCAAATCCCATCTTCTGCACTCAACGCTGTGCTCAGGCTGGGTGGCAGG 1260
Db |||||
QY 1261 GAGCCCTTCCATCTATCTGTGTGTTATCCAGGCTCCAGACTTCTCTCTGCTGC 1320
Db |||||
QY 1261 GAGCCCTTCCATCTATCTGTGTGTTATCCAGGCTCCAGACTTCTCTCTGCTGC 1320
Db |||||
QY 1321 CCAGTCCGCTCCGCTTATCTATCTCTTCTGCTCCGCTCCGCTCCGCT 1380
Db |||||
QY 1381 TCTGTCCCTCCGCTTATCTATCTCTTCTGCTCCGCTCCGCTCCGCTCCGCT 1440
Db |||||
QY 1381 TCTGTCCCTCCGCTTATCTATCTCTTCTGCTCCGCTCCGCTCCGCTCCGCT 1440
Db |||||
QY 1441 GGGCTGGCCAGTGGATTTTCATGGTGATCAATTAAGAAAGAAATCGCAACCAAAAA 1500
Db |||||
QY 1501 AAAAA 1505
Db |||||
QY 1501 AAAAA 1505
Db |||||

RESULT 8
US-10-063-512-127
; Sequence 127, Application US/10063512
; Publication No. US20030018183A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,512
; CURRENT FILING DATE: 2002-05-01
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 127
; LENGTH: 1505
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-063-512-127

Query Match 100.0%; Score 1505; DB 13; Length 1505;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGATCGGACCCAGAGTGGCGCGCGGCGGAGAGCGGCGGCGTCACTCC 60
Db |||||
QY 1 CGCGATCGGACCCAGAGTGGCGCGCGGCGGAGAGCGGCGGCGTCACTCC 60
Db |||||
QY 61 TCGACCCCGGTGCTGGGCTAGTTCAGAGGAGCGGCGGCGGCGGCGGCGGCGGCGG 120
Db |||||
QY 61 TCGACCCCGGTGCTGGGCTAGTTCAGAGGAGCGGCGGCGGCGGCGGCGGCGGCGG 120
Db |||||
QY 121 CCGGATCGGAGCGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
Db |||||
QY 121 CCGGATCGGAGCGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
Db |||||
QY 181 GCGGCGGCGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
Db |||||
QY 181 GCGGCGGCGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
Db |||||
QY 241 ACTGTGGGCAACATCGAGGAGTGGCTGCTGATGTAAGAGTGCAGGCTACCCCGGACT 300
Db |||||

```

```

Db 241 ACTGTGGGCAACATCGAGGAGTGGCTGCTGATGTAAGAGTGCAGGCTACCCCGGACT 300
QY |||||
QY 301 TTGATCCCTTACAGATGTGACCTATCAAAATGAAGAGACATCTCTCTCAATGTTCTAGCT 360
Db |||||
QY 301 TTGATCCCTTACAGATGTGACCTATCAAAATGAAGAGACATCTCTCTCAATGTTCTAGCT 360
Db |||||
QY 361 ATCGTTCTCAGACAGCGGTGTAGACATCTGATCAACATGCTGCTGCTGGCTGGCCGGCT 420
Db |||||
QY 361 ATCGTTCTCAGACAGCGGTGTAGACATCTGATCAACATGCTGCTGCTGGCTGGCCGGCT 420
Db |||||
QY 421 GACACCTCTGCTCTCAGGAGCACACAGTGGTTGGAAGGACATGTTCAATGTGAACGTGCTG 480
Db |||||
QY 421 GACACCTCTGCTCTCAGGAGCACACAGTGGTTGGAAGGACATGTTCAATGTGAACGTGCTG 480
Db |||||
QY 481 GCCCTCAGATCTGCAACCGGAAGCTTACAGTCCATGAAGAGCGGAATGTGAGCAT 540
Db |||||
QY 481 GCCCTCAGATCTGCAACCGGAAGCTTACAGTCCATGAAGAGCGGAATGTGAGCAT 540
Db |||||
QY 541 GGGCAGCATATTAAATAGATGCTGCGCCACCGAGTGTACCCCTGCTGTGACCC 600
Db |||||
QY 541 GGGCAGCATATTAAATAGATGCTGCGCCACCGAGTGTACCCCTGCTGTGACCC 600
Db |||||
QY 601 CACTTCTATAGTGCACCAAGTATGCGCTCACTGCGCTGACAGAGGACTGAGCAAGAG 660
Db |||||
QY 601 CACTTCTATAGTGCACCAAGTATGCGCTCACTGCGCTGACAGAGGACTGAGCAAGAG 660
Db |||||
QY 661 CTTGGAGGCGCCAGACCCATCCAGGACCGAGTGTATCTCACTGCTGCTGAGGAGCA 720
Db |||||
QY 661 CTTGGAGGCGCCAGACCCATCCAGGACCGAGTGTATCTCACTGCTGCTGAGGAGCA 720
Db |||||
QY 721 CAATTGCGCTTCAAACTCCAGCAAGGACCGTGAAGAGGAGTGTATCTCACTGCTGAGGAGCA 780
Db |||||
QY 721 CAATTGCGCTTCAAACTCCAGCAAGGACCGTGAAGAGGAGTGTATCTCACTGCTGAGGAGCA 780
Db |||||
QY 781 ATGAAGTGTCTCAAACTCCAGGAGTGTGCGCGAGGCTGTATCTCACTGCTGAGGAGCA 840
Db |||||
QY 781 ATGAAGTGTCTCAAACTCCAGGAGTGTGCGCGAGGCTGTATCTCACTGCTGAGGAGCA 840
Db |||||
QY 841 GCACATCCAGATTTGAGAGCATCCAGATGAGGCGGCGGAGGAGTGTAGTGTGAGTGTG 900
Db |||||
QY 841 GCACATCCAGATTTGAGAGCATCCAGATGAGGCGGCGGAGGAGTGTAGTGTGAGTGTG 900
Db |||||
QY 901 TGGAGCTCTCTTCCCTCCCACTTCCCACTTCCCACTTCCCACTTCCCACTTCCCACTTCCCACT 960
Db |||||
QY 901 TGGAGCTCTCTTCCCTCCCACTTCCCACTTCCCACTTCCCACTTCCCACTTCCCACTTCCCACT 960
Db |||||
QY 961 TGTGATTTCTGGATCAGGAGTACCACTTCTGCTCCACACCCCGGAGGAGTGTAGGAG 1020
Db |||||
QY 961 TGTGATTTCTGGATCAGGAGTACCACTTCTGCTCCACACCCCGGAGGAGTGTAGGAG 1020
Db |||||
QY 1021 ATTTGTTGAGATTTTATATCATCTGTCMAATGCTTCACTGTCMAATGCTTCACTGTCMAATG 1080
Db |||||
QY 1021 ATTTGTTGAGATTTTATATCATCTGTCMAATGCTTCACTGTCMAATGCTTCACTGTCMAATG 1080
Db |||||
QY 1081 GGCTGGGAGAGGAGTGTGCTCCCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCT 1140
Db |||||
QY 1081 GGCTGGGAGAGGAGTGTGCTCCCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCT 1140
Db |||||
QY 1141 TGGGCACTTGGGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCT 1200
Db |||||
QY 1141 TGGGCACTTGGGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCT 1200
Db |||||
QY 1201 GCCAAATCCCATCTTCTGTCACCTCAACGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
Db |||||
QY 1201 GCCAAATCCCATCTTCTGTCACCTCAACGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
Db |||||
QY 1261 GAGGCTTCACTTATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
Db |||||
QY 1261 GAGGCTTCACTTATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
Db |||||
QY 1321 CCCACTGACCCCTTCTCCCTCTTATCTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
Db |||||

```


Db 1321 CCACCTGACCCCTCTCCCTTATCTATCTCTCTGCTCCCGAGCCAGCTTTGGCT 1380
 Qy 1381 TCTTCTCCCTCTCTGGGTATCCCTCACTCTGACTATGAGTATGGAGAGAACACA 1440
 Db 1381 TCTTCTCCCTCTCTGGGTATCCCTCACTCTGACTATGAGTATGGAGAGAACACA 1440
 Qy 1441 GGGCTGGCCAGTGTATTTATGATGATCAATTAAGAAAGAAATCGAACCAAAAA 1500
 Db 1441 GGGCTGGCCAGTGTATTTATGATGATCAATTAAGAAAGAAATCGAACCAAAAA 1500
 Qy 1501 AAAAA 1505
 Db 1501 AAAAA 1505

RESULT 9

US-10-063-513-127
 ; Sequence 127, Application US/10063513
 ; Publication No. US20030018172A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baton, Dan L.
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, Christopher J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE OF INVENTION: ACIDS ENCODING THE SAME
 ; FILE REFERENCE: P3230R1C1
 ; CURRENT APPLICATION NUMBER: US/10/063,513
 ; CURRENT FILING DATE: 2002-05-01
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 170
 ; SEQ ID NO 127
 ; LENGTH: 1505
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 US-10-063-513-127

Query Match 100.0%; Score 1505; DB 13; Length 1505;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 CGCGGATCGGACCCAGCAGTTCGGCGCGCGGCGGAGAGAGCGCGCGGTCTGAGTCC 60
 Db 1 CGCGGATCGGACCCAGCAGTTCGGCGCGCGGCGGAGAGAGCGCGCGGTCTGAGTCC 60
 Qy 61 TCGACCCCGGTGTGGGTAGTTCAGCGAGCGCGGCGGCGGTGGGCGGCGGCGGCGG 120
 Db 61 TCGACCCCGGTGTGGGTAGTTCAGCGAGCGCGGCGGCGGTGGGCGGCGGCGGCGG 120
 Qy 121 CCGGCGATGGAGCGGTGGCGGACCGGTGCGGTGAGCGGGGCGGCGGCGGCGGCGATC 180
 Db 121 CCGGCGATGGAGCGGTGGCGGACCGGTGCGGTGAGCGGGGCGGCGGCGGCGGCGATC 180
 Qy 181 GCGCGCGCGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
 Db 181 GCGCGCGCGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
 Qy 241 ACTGTGGGCAACATCGAGGAGTGGCTGCTGAATGTAAAGTGCAGGCTACCCCGGACT 300
 Db 241 ACTGTGGGCAACATCGAGGAGTGGCTGCTGAATGTAAAGTGCAGGCTACCCCGGACT 300
 Qy 301 TTGATCCCTACAGATGACCTATCAATGAAGAGGACATCCCTCTCCATGTTCTCAGCT 360
 Db 301 TTGATCCCTACAGATGACCTATCAATGAAGAGGACATCCCTCTCCATGTTCTCAGCT 360
 Qy 361 ATCCGTTCTCAGCAGCGGTGACATCTGCAATGTCACATGTCGCTTGGCGCGGCGCT 420
 Db 361 ATCCGTTCTCAGCAGCGGTGACATCTGCAATGTCACATGTCGCTTGGCGCGGCGCT 420

Qy 421 GACACCTCTCTCAGGAGAGCACAGTGGTGAAGAGGACATGTTCAATGTGAACGTGCTG 480
 Db 421 GACACCTCTCTCAGGAGAGCACAGTGGTGAAGAGGACATGTTCAATGTGAACGTGCTG 480
 Qy 481 GCCCTCAGCATCTGCACACGGGAGGCTTACAGTCCATGAAGGAGGGAATGTGACCAT 540
 Db 481 GCCCTCAGCATCTGCACACGGGAGGCTTACAGTCCATGAAGGAGGGAATGTGACCAT 540
 Qy 541 GGGCACATCAATTAACATAGCATGTCTGGCCACCGAGTGTACCCCTGTCTGTGACC 600
 Db 541 GGGCACATCAATTAACATAGCATGTCTGGCCACCGAGTGTACCCCTGTCTGTGACC 600
 Qy 601 CACTTCTATAGTGGCCACCAAGTATGCCCTCATCTGGCTGACAGAGGACATGAGGCAAG 660
 Db 601 CACTTCTATAGTGGCCACCAAGTATGCCCTCATCTGGCTGACAGAGGACATGAGGCAAG 660
 Qy 661 CTTGGGAGGCGCCAGAGCCACATCCGAGCCAGTGCATCTCTCCAGGTGTGGTGAGACA 720
 Db 661 CTTGGGAGGCGCCAGAGCCACATCCGAGCCAGTGCATCTCTCCAGGTGTGGTGAGACA 720
 Qy 721 CAATTGCGCTTCAAACTCCAGCAAGACCTTGAGAAAGGAGCTGCCACCTATGAGCAA 780
 Db 721 CAATTGCGCTTCAAACTCCAGCAAGACCTTGAGAAAGGAGCTGCCACCTATGAGCAA 780
 Qy 781 ATGAAGTGTCTCAAACTCCAGCAAGACCTTGAGAAAGGAGCTGCCACCTATGAGCAA 840
 Db 781 ATGAAGTGTCTCAAACTCCAGCAAGACCTTGAGAAAGGAGCTGCCACCTATGAGCAA 840
 Qy 841 GCACATCCAGATTGGAGACATCCAGATGAGGCCACGAGGAGGAGCTGACCTAGTACTG 900
 Db 841 GCACATCCAGATTGGAGACATCCAGATGAGGCCACGAGGAGGAGCTGACCTAGTACTG 900
 Qy 901 TGGAGGCTCTCTCCCTCCCGAGGATGTGGCGAGGCTGTATCTACGTCTCAGCACCCCC 960
 Db 901 TGGAGGCTCTCTCCCTCCCGAGGATGTGGCGAGGCTGTATCTACGTCTCAGCACCCCC 960
 Qy 961 TGTGATTTCTGGATCAGGAGTACCATCTCTGTCACACCCCGAGGAGCTGAGAAA 1020
 Db 961 TGTGATTTCTGGATCAGGAGTACCATCTCTGTCACACCCCGAGGAGCTGAGAAA 1020
 Qy 1021 ATTTGTTTGAGATTTTATATCATCTTGTCAAAATGCTTCAATTTGTAATGTGAATAATG 1080
 Db 1021 ATTTGTTTGAGATTTTATATCATCTTGTCAAAATGCTTCAATTTGTAATGTGAATAATG 1080
 Qy 1081 GGCTGGGAAAGAGGAGTGGTCCCTAAATGTTTAACTTGTCTTGTGCCCC 1140
 Db 1081 GGCTGGGAAAGAGGAGTGGTCCCTAAATGTTTAACTTGTCTTGTGCCCC 1140
 Qy 1141 TGGGCACTTGGCCCTTCTGCTCTCAGTGTCTTCCCTTTGACATGGGAAAGGAGTGTG 1200
 Db 1141 TGGGCACTTGGCCCTTCTGCTCTCAGTGTCTTCCCTTTGACATGGGAAAGGAGTGTG 1200
 Qy 1201 GCCAAAATCCCATCTTCTTGGACCTCAAGCTCTGTGGCTCAGGCTGGGGTGGCAGAGG 1260
 Db 1201 GCCAAAATCCCATCTTCTTGGACCTCAAGCTCTGTGGCTCAGGCTGGGGTGGCAGAGG 1260
 Qy 1261 GAGGCTTCACTTATATCTGTGTGTATCCAGGCTTCCAGACTTCTCTCTGCTGCTG 1320
 Db 1261 GAGGCTTCACTTATATCTGTGTGTATCCAGGCTTCCAGACTTCTCTCTGCTGCTG 1320
 Qy 1321 CCACCTGACCTCTCCCTTATCTATCTCTGCTTCCCTCCCGAGGAGTCTTGGCT 1380
 Db 1321 CCACCTGACCTCTCCCTTATCTATCTCTGCTTCCCTCCCGAGGAGTCTTGGCT 1380
 Qy 1381 TCTTGTCCCTCTCTGGGCTCATCCCTCAGCTCTGACTCTGACTATGGCAGCAGAACCA 1440
 Db 1381 TCTTGTCCCTCTCTGGGCTCATCCCTCAGCTCTGACTCTGACTATGGCAGCAGAACCA 1440
 Qy 1441 GGGCTGGCCAGTGGATTTTCATGGTGATCAATTAAGAAAGAAATCGAACCAAAAA 1500
 Db 1441 GGGCTGGCCAGTGGATTTTCATGGTGATCAATTAAGAAAGAAATCGAACCAAAAA 1500

APPLICANT: Eaton, Dan L.
APPLICANT: Filvaroff, Ellen
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3230R1C1
CURRENT APPLICATION NUMBER: US/10/063,551
CURRENT FILING DATE: 2002-05-02
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 170
SEQ ID NO 127
LENGTH: 1505
TYPE: DNA
ORGANISM: Homo Sapien
US-10-063-551-127

Query Match 100.0%; Score 1505; DB 13; Length 1505;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGCATCGGACCCCAAGCAGGTGCGGGCGGCGGAGAGCGCGCGGGGCGTCAAGTCC 60
Db 1 CGCGCATCGGACCCCAAGCAGGTGCGGGCGGCGGAGAGCGCGCGGGGCGTCAAGTCC 60

Qy 61 TCGACCCCGGTGTCGGGTAGTCCAGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
Db 61 TCGACCCCGGTGTCGGGTAGTCCAGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120

Qy 121 CCCGCATGGAGCGGTGCGGCGACCGCGTGGCGTGGTGA CGGGGGCGCTCGGGGGGCGATC 180
Db 121 CCCGCATGGAGCGGTGCGGCGACCGCGTGGCGTGGTGA CGGGGGCGCTCGGGGGGCGATC 180

Qy 181 GCGCGGCGGTGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
Db -81 GCGCGGCGGTGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240

Qy 241 ACTGTGGCAACATCGAGGAGTGGCTGCTGAAATGTAAGAGTGCAGGCTACCCCGGGACT 300
Db 241 ACTGTGGCAACATCGAGGAGTGGCTGCTGAAATGTAAGAGTGCAGGCTACCCCGGGACT 300

Qy 301 TTGATCCCTACAGATGACCTATCAAAATGAAGAGGACATCCTCTCCATGTTCTCAGCT 360
Db 301 TTGATCCCTACAGATGACCTATCAAAATGAAGAGGACATCCTCTCCATGTTCTCAGCT 360

Qy 361 ATCCGTTCTCAGCAGCGGTGACATCTGCATCAACATGCTGGCTTGGCCCGGCGCT 420
Db 361 ATCCGTTCTCAGCAGCGGTGACATCTGCATCAACATGCTGGCTTGGCCCGGCGCT 420

Qy 421 GACACCTGCTCTCAGGCAGCACAGTGGTGGAGGACATGTTCAATGTGAACGTGCTG 480
Db 421 GACACCTGCTCTCAGGCAGCACAGTGGTGGAGGACATGTTCAATGTGAACGTGCTG 480

Qy 481 GCCCTCAGCATCTGCACAGGAGGCGCTACAGTCCATGAAGAGCGGAATGTGGAGAT 540
Db 481 GCCCTCAGCATCTGCACAGGAGGCGCTACAGTCCATGAAGAGCGGAATGTGGAGAT 540

Qy 541 GGGCACATCAATTAACATAGCATGCTGGCCACCGAGTGTACCCCTGTCTGTGACC 600
Db 541 GGGCACATCAATTAACATAGCATGCTGGCCACCGAGTGTACCCCTGTCTGTGACC 600

Qy 601 CACTTCTATAGTGCACCAAGATGTCGGTCACTCGCTGACAGAGGACTGAGCAAGAG 660
Db 601 CACTTCTATAGTGCACCAAGATGTCGGTCACTCGCTGACAGAGGACTGAGCAAGAG 660

Qy 661 CTTGGGAGGCGGCGGAGCCCATCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
Db 661 CTTGGGAGGCGGCGGAGCCCATCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720

Qy 721 CAATTGCGCTTCAAACTCCACGACGAGGAGCCCTGAGAGGCGAGCTGCCACTATGAGCAA 780
Db 721 CAATTGCGCTTCAAACTCCACGACGAGGAGCCCTGAGAGGCGAGCTGCCACTATGAGCAA 780

Qy 781 ATGAAGTGTCTCAAAACCCGAGGATGTGCCGAGGCTGTATCTAGTCTCTAGCAGCCCC 840
Db 781 ATGAAGTGTCTCAAAACCCGAGGATGTGCCGAGGCTGTATCTAGTCTCTAGCAGCCCC 840

Qy 841 GCACACATCCAGATTGGAGACATCCAGATGAGGCCACCGAGAGCGGTGACCTAGTGACTG 900
Db 841 GCACACATCCAGATTGGAGACATCCAGATGAGGCCACCGAGAGCGGTGACCTAGTGACTG 900

Qy 901 TGGGAGCTCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 960
Db 901 TGGGAGCTCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 960

Qy 961 TGTGATTTCTGGATCAGGGATACCACTTCTCTGTCACACCCGAGCGGCTAGAAA 1020
Db 961 TGTGATTTCTGGATCAGGGATACCACTTCTCTGTCACACCCGAGCGGCTAGAAA 1020

Qy 1021 ATTGTTTGAGATTTTATATCATCTTGTCAAATTTGCTTCAGTTGTAAATGTGAAAAATG 1080
Db 1021 ATTGTTTGAGATTTTATATCATCTTGTCAAATTTGCTTCAGTTGTAAATGTGAAAAATG 1080

Qy 1081 GGCTGGGAAAAAGGAGGTGTCCTTAATTTTACTTTTAACTTTCTTTGTCGCC 1140
Db 1081 GGCTGGGAAAAAGGAGGTGTCCTTAATTTTACTTTTAACTTTCTTTGTCGCC 1140

Qy 1141 TGGGCACTTGGCCCTTTGTCGCTCTCAGTGTCCTTCCCTTTGACATGGGAAAGGAGTGTG 1200
Db 1141 TGGGCACTTGGCCCTTTGTCGCTCTCAGTGTCCTTCCCTTTGACATGGGAAAGGAGTGTG 1200

Qy 1201 GCCAAATCCCACTTCTTTCACCTCAACGCTCTGTGGCTCAGGGCTGGGCTGGCAGAGG 1260
Db 1201 GCCAAATCCCACTTCTTTCACCTCAACGCTCTGTGGCTCAGGGCTGGGCTGGCAGAGG 1260

Qy 1261 GAGGCTTCACTTATATCTGTGTGTTATCCAGGGCTCCAGACTTCTCTCTGCTGCTGC 1320
Db 1261 GAGGCTTCACTTATATCTGTGTGTTATCCAGGGCTCCAGACTTCTCTCTGCTGCTGC 1320

Qy 1321 CCCACTGACCTCTCCCTCTTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1380
Db 1321 CCCACTGACCTCTCCCTCTTATCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1380

Qy 1381 TCTTGTCCCTCTCTGGGGTCACTCCCTCCACTCTGACTCTGACTATGGCAGCAGAACCA 1440
Db 1381 TCTTGTCCCTCTCTGGGGTCACTCCCTCCACTCTGACTCTGACTATGGCAGCAGAACCA 1440

Qy 1441 GGGCTGGCCAGTGGATTTTCATGGTGTATTAATAAAGAGAAATCGAACCAAAAAA 1500
Db 1441 GGGCTGGCCAGTGGATTTTCATGGTGTATTAATAAAGAGAAATCGAACCAAAAAA 1500

Qy 1501 AAAAA 1505
Db 1501 AAAAA 1505

RESULT 12
US-10-174-581-377
Sequence 377, Application US/10174581
Publication No. US20030017540A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C41
CURRENT APPLICATION NUMBER: US/10/174,581
CURRENT FILING DATE: 2002-06-18
PRIOR APPLICATION NUMBER: 10/052586
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063540
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063564
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063734
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063870
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066120
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066466
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066772
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069335
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069425
PRIOR FILING DATE: 1997-12-12
PRIOR APPLICATION NUMBER: 60/069870
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/068017
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081049

PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086023
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/086392
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/086486
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087098
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087208
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05

```

; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088722
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088740
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088811
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088825
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088863
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089090
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653

Query Match      100.0%; Score 1505; DB 13; Length 1505;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGGATCGGACCCAGCAGGTGCGCGGGCGGCGGAGAGCGCGCGGCGGTCAAGTCC 60
DB 1 CGCGGATCGGACCCAGCAGGTGCGCGGGCGGCGGAGAGCGCGCGGCGGTCAAGTCC 60

QY 61 TCGACCCCGGTGTCGGGCTAGTCCAGCGAGCGGCGGCGGCGGTGGGCCCATGGCCAGG 120
DB 61 TCGACCCCGGTGTCGGGCTAGTCCAGCGAGCGGCGGCGGCGGTGGGCCCATGGCCAGG 120

QY 121 CCCGCGATGAGCGGTGCGCGACCGGTGCGGCTGGCGTGGCGTGGCGGCGCTCGGGGGCATC 180
DB 121 CCCGCGATGAGCGGTGCGCGACCGGTGCGGCTGGCGTGGCGTGGCGGCGCTCGGGGGCATC 180

QY 181 GCGCGGCGGTGGCCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
DB 181 GCGCGGCGGTGGCCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240

QY 241 ACTGTGGGCAACATCGAGGAGTGGCTGCTGAATGTAAGTGCAGGCTACCCGGGACT 300
DB 241 ACTGTGGGCAACATCGAGGAGTGGCTGCTGAATGTAAGTGCAGGCTACCCGGGACT 300

QY 301 TTGATCCCTTACAGATGACCTATCAATGAAGGACATCCCTCTCCATGTTCTCAGCT 360
DB 301 TTGATCCCTTACAGATGACCTATCAATGAAGGACATCCCTCTCCATGTTCTCAGCT 360

QY 361 ATCCGTTCTACGACAGCGGTGAGACATCTGCATCAACATGCTGGCTTGGCCGGGCT 420
DB 361 ATCCGTTCTACGACAGCGGTGAGACATCTGCATCAACATGCTGGCTTGGCCGGGCT 420

```

```

QY 421 GACACCCCTGCTCTCAGGAGACACAGTGGTTGGAAGGACATGTTCAATGTGAACGTCGCTG 480
DB 421 GACACCCCTGCTCTCAGGAGACACAGTGGTTGGAAGGACATGTTCAATGTGAACGTCGCTG 480

QY 481 GCGCTCAGCATCTGCACACGGGAAGCCTACAGTGCATGAAGGAGCGGAATGTGGAGGAT 540
DB 481 GCGCTCAGCATCTGCACACGGGAAGCCTACAGTGCATGAAGGAGCGGAATGTGGAGGAT 540

QY 541 GGGCACATCATTAACATCAATAGCATGTCTGGCCACCCGAGTGTACCCCTGCTGTGTGACC 600
DB 541 GGGCACATCATTAACATCAATAGCATGTCTGGCCACCCGAGTGTACCCCTGCTGTGTGACC 600

QY 601 CACTTCTATAGTGCACACCAAGTATGCGGTCACTGGCTGCACAGAGGACTGAGGCAAGAG 660
DB 601 CACTTCTATAGTGCACACCAAGTATGCGGTCACTGGCTGCACAGAGGACTGAGGCAAGAG 660

QY 661 CTTGGGAGGCGGACAGACCCACATCCGAGCCAGTGCATCTCTCCAGGTCGTGGTGAGACA 720
DB 661 CTTGGGAGGCGGACAGACCCACATCCGAGCCAGTGCATCTCTCCAGGTCGTGGTGAGACA 720

QY 721 CAATTCGCTTCAAACTCCAGCAAGGACCCCTGAGAAAGGCGAGCTGCCACCTATGAGCAA 780
DB 721 CAATTCGCTTCAAACTCCAGCAAGGACCCCTGAGAAAGGCGAGCTGCCACCTATGAGCAA 780

QY 781 ATGAAGTGTCTCAAAACCCGAGGATGTGGCGAGGCTGTATCTACGTCTCAGCACCCGCC 840
DB 781 ATGAAGTGTCTCAAAACCCGAGGATGTGGCGAGGCTGTATCTACGTCTCAGCACCCGCC 840

QY 841 GCACACATCCAGATTGGAGACATCCAGATGAGGCCACGAGGAGCTGACCTAGTGTGACTG 900
DB 841 GCACACATCCAGATTGGAGACATCCAGATGAGGCCACGAGGAGCTGACCTAGTGTGACTG 900

QY 901 TGGGAGCTCTCTCTCCCTCCCACTTCCAGTGGCTTGCCTCTCTCTCTCTCTCTCTCTCTCT 960
DB 901 TGGGAGCTCTCTCTCTCCCTCCCACTTCCAGTGGCTTGCCTCTCTCTCTCTCTCTCTCTCT 960

QY 961 TGTGATTTCTGGATCAGGGATACCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1020
DB 961 TGTGATTTCTGGATCAGGGATACCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1020

QY 1021 ATTTGTTTGGATTTTATATATATATATATATATATATATATATATATATATATATATAT 1080
DB 1021 ATTTGTTTGGATTTTATATATATATATATATATATATATATATATATATATATATATAT 1080

QY 1081 GCGTGGGAAAAGGAGGTGTTCCCTAAATGTTTAACTTGTAACTTCTCTCTCTCTCTCTCTCT 1140
DB 1081 GCGTGGGAAAAGGAGGTGTTCCCTAAATGTTTAACTTGTAACTTCTCTCTCTCTCTCTCT 1140

QY 1141 TGGGCACTTGGGCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1200
DB 1141 TGGGCACTTGGGCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1200

QY 1201 GCCAAAATCCCACTCTCTTCCACCTCAACGCTGTGTGGCTCAGGCTCGGGTGGCAGAGG 1260
DB 1201 GCCAAAATCCCACTCTCTTCCACCTCAACGCTGTGTGGCTCAGGCTCGGGTGGCAGAGG 1260

QY 1261 GAGGCTTTCACCTTATATATCTGTGTGTGTATCCAGGGCTCCAGACTTCTCTCTCTCTCTCT 1320
DB 1261 GAGGCTTTCACCTTATATATCTGTGTGTGTATCCAGGGCTCCAGACTTCTCTCTCTCTCT 1320

QY 1321 CCCACTGACCCCTCTCCCTTATCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1380
DB 1321 CCCACTGACCCCTCTCCCTTATCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1380

QY 1381 TCTTGTCCCTCTCTGGGCTCATCCCTCCACTCTGACTCTGACTATGGCAGGAGCAACACCA 1440
DB 1381 TCTTGTCCCTCTCTGGGCTCATCCCTCCACTCTGACTCTGACTATGGCAGGAGCAACACCA 1440

QY 1441 GGGCTTGGCCAGTGGATTTTCAATGGTGATATTTAAAAAGAAAAATCGCAACCAAAAAA 1500
DB 1441 GGGCTTGGCCAGTGGATTTTCAATGGTGATATTTAAAAAGAAAAATCGCAACCAAAAAA 1500

```



```

RESULT 15
US-10-176-914-377
; Sequence 377, Application US/10176914
; Publication No. US20030017543A1
; GENERAL INFORMATION:
; APPLICANT: Baker Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.

```



```

; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P430R1C83
; CURRENT APPLICATION NUMBER: US/10/176,914
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 377
; LENGTH: 1505
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-914-377

Query Match      100.0%; Score 1505; DB 13; Length 1505;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  CGCGATCGGACCCCAAGCAGGTGCGCGCGCGCGGAGAGAGCGCGCGCGGTCTCAGCTCC 60
DB      1  CGCGATCGGACCCCAAGCAGGTGCGCGCGCGCGGAGAGAGCGCGCGCGGTCTCAGCTCC 60
QY      61  TCGACCCCGGTGTGCGGTGTGTCAGCGAGCGGAGCGGCGGCGGTGCGGCGGCGGCGGCGG 120
DB      61  TCGACCCCGGTGTGCGGTGTGTCAGCGAGCGGAGCGGCGGCGGTGCGGCGGCGGCGGCGG 120
QY      121  CCGGCGATGAGCGGTGCGGCGGCGGCGGTGCGGTGCGGTGCGGCGGCGGTGCGGCGGCGGCG 180
DB      121  CCGGCGATGAGCGGTGCGGCGGCGGCGGTGCGGTGCGGTGCGGCGGCGGTGCGGCGGCGGCG 180
QY      181  GCGCGCGCGGTGCGGCGGCGGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTG 240
DB      181  GCGCGCGCGGTGCGGCGGCGGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTG 240
QY      241  ACTGTGGGCAACATCGAGGAGTGTGCTGTAATGTAAGTGAAGTGAAGTGAAGTGAAGTGAAG 300
DB      241  ACTGTGGGCAACATCGAGGAGTGTGCTGTAATGTAAGTGAAGTGAAGTGAAGTGAAGTGAAG 300
QY      301  TTGATCCCTTACAGATGTGACCTATCAATGAAGAGGACATCTCCATGTTCTCAGCT 360
DB      301  TTGATCCCTTACAGATGTGACCTATCAATGAAGAGGACATCTCCATGTTCTCAGCT 360
QY      361  ATCCGTTCTCAGCAGCGGTTAGACATCTGCATCAACAATGCTGGCTTGGCCCGGCT 420
DB      361  ATCCGTTCTCAGCAGCGGTTAGACATCTGCATCAACAATGCTGGCTTGGCCCGGCT 420
QY      421  GACACCTGCTCTCAGGAGCAACAGTGTGGAAGGACATGTTCAATGTGAACGTGTG 480
DB      421  GACACCTGCTCTCAGGAGCAACAGTGTGGAAGGACATGTTCAATGTGAACGTGTG 480
QY      481  GCGCTCAGCATCTGACAGCGGAGCGCTACAGTCCATGAAGAGCGGAACTGGACCAT 540
DB      481  GCGCTCAGCATCTGACAGCGGAGCGCTACAGTCCATGAAGAGCGGAACTGGACCAT 540
QY      541  GGGCACATCAATTAACATCAATAGCATGTCTGGCCACCGAGTGTATCCCGTGTGTGACC 600
DB      541  GGGCACATCAATTAACATCAATAGCATGTCTGGCCACCGAGTGTATCCCGTGTGTGACC 600
QY      601  CACTTCTATAGTGCCCAACCAAGTATGCGTCACTGCGCTGACAGAGGAACTGAGGCAAGAG 660
DB      601  CACTTCTATAGTGCCCAACCAAGTATGCGTCACTGCGCTGACAGAGGAACTGAGGCAAGAG 660
QY      661  CTTGGGAGGCGGAGCCACATCCAGCCAGTGTCTCTCCAGGTGTGGTGGAGCA 720
DB      661  CTTGGGAGGCGGAGCCACATCCAGCCAGTGTCTCTCCAGGTGTGGTGGAGCA 720
QY      721  CAATTGCGCTTCAAACTCCAGGACAGGACCTCGAGAGGCGAGTGTCCACCTATGAGCAA 780
DB      721  CAATTGCGCTTCAAACTCCAGGACAGGACCTCGAGAGGCGAGTGTCCACCTATGAGCAA 780

```

```

QY      781  ATGAAGTGTCTCAAAACCGAGGATGTGCGCGAGGCTGTTATCTACGTCTCAGCACCCCC 840
DB      781  ATGAAGTGTCTCAAAACCGAGGATGTGCGCGAGGCTGTTATCTACGTCTCAGCACCCCC 840
QY      841  GCACACATCCAGATTGGAGACATCCAGATGAGGCGCCACGAGCAGGTGACCTAGTACTG 900
DB      841  GCACACATCCAGATTGGAGACATCCAGATGAGGCGCCACGAGCAGGTGACCTAGTACTG 900
QY      901  TGGAGGCTCTCTTCCCTCCCAACCTTCAATGCTTCCCTCTGCTCTGCTCTGCTCTGCT 960
DB      901  TGGAGGCTCTCTTCCCTCCCAACCTTCAATGCTTCCCTCTGCTCTGCTCTGCTCTGCT 960
QY      961  TGTGATTTCTGGATCAGGGATACCACTTCTCTGTCACACCCCGACCGAGGGCTAGAAA 1020
DB      961  TGTGATTTCTGGATCAGGGATACCACTTCTCTGTCACACCCCGACCGAGGGCTAGAAA 1020
QY      1021  ATTTGTTTGAGATTTTATATCATCTTGTCAAATTTGCTTCAGTTGTAATGTGAAAAATG 1080
DB      1021  ATTTGTTTGAGATTTTATATCATCTTGTCAAATTTGCTTCAGTTGTAATGTGAAAAATG 1080
QY      1081  GGCTGGGGAAGAGGAGGTGCTCCCTAATGTTTTTACTTGTAACTTGTCTTGTGCCCC 1140
DB      1081  GGCTGGGGAAGAGGAGGTGCTCCCTAATGTTTTTACTTGTAACTTGTCTTGTGCCCC 1140
QY      1141  TGGGCACTTGGGCTTGTCTGCTCTCAGTGTCTTCCCTTTGACATGGGAAAGAGTTGTG 1200
DB      1141  TGGGCACTTGGGCTTGTCTGCTCTCAGTGTCTTCCCTTTGACATGGGAAAGAGTTGTG 1200
QY      1201  GCCAAATCCCATCTTCTTGCACTCAACGCTCTGTGGCTCAGGGCTGGGGTGGCAGAGG 1260
DB      1201  GCCAAATCCCATCTTCTTGCACTCAACGCTCTGTGGCTCAGGGCTGGGGTGGCAGAGG 1260
QY      1261  GAGGCTTCACTTATATCTGTGTTGTTATCCAGGGCTCCAGACTTCTCTCTCTGCTGC 1320
DB      1261  GAGGCTTCACTTATATCTGTGTTGTTATCCAGGGCTCCAGACTTCTCTCTCTGCTGC 1320
QY      1321  CCCACTGCACTCTCCCTTATCTATCTCTCTCTGCTCTCCAGGCTCCAGGCTCTGCT 1380
DB      1321  CCCACTGCACTCTCCCTTATCTATCTCTCTCTGCTCTCCAGGCTCCAGGCTCTGCT 1380
QY      1381  TCTTGTCCCTCTCTGGGGTCTATCCCTCCACTCTGACTCTGACTATGGCAGCAGAACCA 1440
DB      1381  TCTTGTCCCTCTCTGGGGTCTATCCCTCCACTCTGACTCTGACTATGGCAGCAGAACCA 1440
QY      1441  GGGCTTGGCCAGTGGAATTTCAATGATGATCAATTAAGGAAAGAAATCGCAACCAAAAA 1500
DB      1441  GGGCTTGGCCAGTGGAATTTCAATGATGATCAATTAAGGAAAGAAATCGCAACCAAAAA 1500
QY      1501  AAAAA 1505
DB      1501  AAAAA 1505

```

Search completed: September 16, 2004, 16:48:49
Job time : 743 secs

Blank

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2004, 12:44:01 ; Search time 600 Seconds
(without alignments)
10655.897 Million cell updates/sec

Title: US-10-063-735-127

Perfect score: 1505

Sequence: 1 cggcgatcgaccacagcag.....tcgcaaccccccaaaaaa 1505

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 337363 seqs, 212409041 residues

Word size : 10

Total number of hits satisfying chosen parameters: 2278698

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : N Geneseq_29Jan04.*

1: Geneseq1980s.*

2: Geneseq1990s.*

3: Geneseq2000s.*

4: Geneseq2001as.*

5: Geneseq2001bs.*

6: Geneseq2002s.*

7: Geneseq2003as.*

8: Geneseq2003bs.*

9: Geneseq2003cs.*

10: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1505	100.0	1505	4	AAS46113 Human PRO
2	1505	100.0	1505	4	Aaf92121 Human PRO
3	1505	100.0	1505	6	ABs74441 Human CDN
4	1505	100.0	1505	7	ABx78716 Human PRO
5	1505	100.0	1505	7	ACA75688 Novel hum
6	1505	100.0	1505	7	ACA71168 Human sec
7	1505	100.0	1505	7	ACC87696 Human sec
8	1505	100.0	1505	7	ACC87082 Human sec
9	1505	100.0	1505	7	ACD04255 Human sec
10	1505	100.0	1505	7	ACA69586 cDNA enco
11	1505	100.0	1505	7	ACA90431 Novel hum
12	1505	100.0	1505	7	ACC89538 Human sec
13	1505	100.0	1505	7	ACA98329 Novel hum
14	1505	100.0	1505	7	ACA93971 Human sec
15	1505	100.0	1505	7	ACD15364 Human sec
16	1505	100.0	1505	7	ACD08951 Human sec
17	1505	100.0	1505	7	ACC96871 Human sec
18	1505	100.0	1505	7	ACF15592 Human sec
19	1505	100.0	1505	7	ACA72959 Human PRO
20	1505	100.0	1505	7	ACD03131 Novel hum
21	1505	100.0	1505	7	ACD01946 Novel hum
22	1505	100.0	1505	7	ACA92138 Novel hum
23	1505	100.0	1505	7	ACA89563 cDNA enco

24	1505	100.0	1505	7	ACA73573	Human sec
25	1505	100.0	1505	7	ACA05888	Human sec
26	1505	100.0	1505	7	ACA66722	CDNA enco
27	1505	100.0	1505	7	ACA91227	Novel hum
28	1505	100.0	1505	7	ACD81604	Human CDN
29	1505	100.0	1505	7	ACF20297	Human sec
30	1505	100.0	1505	7	ACF19683	Human sec
31	1505	100.0	1505	7	ACD21971	Human sec
32	1505	100.0	1505	7	ACF13136	Human sec
33	1505	100.0	1505	7	ACD25239	Human sec
34	1505	100.0	1505	7	ACF00288	Human sec
35	1505	100.0	1505	7	ACA60426	Novel hum
36	1505	100.0	1505	7	ACA72345	Novel hum
37	1505	100.0	1505	7	ACD04869	Novel hum
38	1505	100.0	1505	7	ACD18330	Human sec
39	1505	100.0	1505	7	ACD08337	Human sec
40	1505	100.0	1505	7	ACA88771	Novel hum
41	1505	100.0	1505	7	ACA70213	Human sec
42	1505	100.0	1505	7	ACD12435	Novel hum
43	1505	100.0	1505	7	ACC74350	Human sec
44	1505	100.0	1505	7	ACD15978	Human sec
45	1505	100.0	1505	7	ACD25546	Novel hum

ALIGNMENTS

RESULT 1
AAS46113
ID AAS46113 standard; cDNA; 1505 BP.
XX
AC AAS46113;
XX

DT 18-DEC-2001 (first entry)
XX

DE Human DNA encoding PRO polypeptide sequence #189.
XX

KW PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep; ss;
KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
KW adrenal; lung; breast; prostate; cervix; liver; genetic disorder;
KW PCR primer.
XX

OS Homo sapiens.
XX

PN WO200168848-A2.
XX

PD 20-SEP-2001.
XX

PF 28-FEB-2001; 2001WO-US006520.
XX

PR 01-MAR-2000; 2000WO-US005601.
PR

PR 02-MAR-2000; 2000WO-US005841.
PR

PR 03-MAR-2000; 2000US-0187202P.
PR

PR 06-MAR-2000; 2000US-0186968P.
PR

PR 14-MAR-2000; 2000US-0189320P.
PR

PR 14-MAR-2000; 2000US-0189328P.
PR

PR 15-MAR-2000; 2000WO-US006884.
PR

PR 21-MAR-2000; 2000US-0190828P.
PR

PR 21-MAR-2000; 2000US-0191007P.
PR

PR 21-MAR-2000; 2000US-0191048P.
PR

PR 28-MAR-2000; 2000US-0191314P.
PR

PR 29-MAR-2000; 2000US-0190302P.
PR

PR 29-MAR-2000; 2000US-0193053P.
PR

PR 30-MAR-2000; 2000WO-US008439.
PR

PR 04-APR-2000; 2000US-0194449P.
PR

PR 04-APR-2000; 2000US-0194647P.
PR

PR 11-APR-2000; 2000US-0195975P.
PR

PR 11-APR-2000; 2000US-0196000P.
PR

PR 11-APR-2000; 2000US-0196187P.
PR

PR 11-APR-2000; 2000US-0196690P.
PR

PR 11-APR-2000; 2000US-0196820P.
PR

PR 18-APR-2000; 2000US-0198121P.
 PR 18-APR-2000; 2000US-0198585P.
 PR 25-APR-2000; 2000US-0199397P.
 PR 25-APR-2000; 2000US-0199550P.
 PR 25-APR-2000; 2000US-0199654P.
 PR 03-MAY-2000; 2000US-0201516P.
 PR 17-MAY-2000; 2000WO-US013705.
 PR 22-MAY-2000; 2000WO-US014042.
 PR 30-MAY-2000; 2000WO-US014941.
 PR 02-JUN-2000; 2000WO-US015264.
 PR 05-JUN-2000; 2000US-0209832P.
 PR 28-JUL-2000; 2000WO-US020710.
 PR 22-AUG-2000; 2000US-00644848.
 PR 24-AUG-2000; 2000WO-US023328.
 PR 08-NOV-2000; 2000WO-US030952.
 PR 01-DEC-2000; 2000WO-US032678.
 PR 20-DEC-2000; 2000WO-US034956.
 XX (GETH) GENENTECH INC.
 PA
 XX Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
 PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
 XX P-PSDB; AAU29212.
 DR WPI; 2001-602746/68.
 XX
 PT Novel nucleic acids encoding PRO polypeptides, used to diagnose the
 PT presence of tumors, such as prostate and breast tumors, in mammals and to
 PT screen for modulators of the compounds.
 XX
 PS Claim 2; Fig 377; 774pp; English.
 XX
 PS Sequences AAS45925-AAS46231 represent DNA molecules encoding and PCR
 CC primers for PRO polypeptides of the invention. The sequences of the
 CC invention can be used to detect the presence of a tumour in a mammal by
 CC comparing the level of expression of a PRO polypeptide in a test sample
 CC of cells from the animal and a control sample of normal cells, whereby a
 CC higher level of expression in the test sample indicates the presence of a
 CC tumour in the mammal. Mammals include dogs, cats, cattle, horses, sheep,
 CC pigs, goats and rabbits but are preferably human. The polypeptides can be
 CC used to stimulate tumour necrosis factor (TNF) alpha release from human
 CC blood, when contacted with it. A specific polypeptide can be used to
 CC stimulate the proliferation or differentiation of chondrocyte cells. The
 CC PRO proteins can be used to determine the presence of tumours and also
 CC susceptibility to tumour development, particularly adrenal, lung, colon,
 CC breast, prostate, rectal, cervical, or liver tumours, in mammalian
 CC subjects. The oligonucleotide probes specific for the PRO nucleic acids
 CC can be used for genetic analysis of individuals with genetic disorders
 XX
 SQ Sequence 1505 BP; 309 A; 442 C; 419 G; 335 T; 0 U; 0 Other;
 Query Match 100.0%; Score 1505; DB 4; Length 1505;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGCGATCGGACCCAAAGCAGGTGCGCGCGCGGCGGAGAGCGCGCGGCGGTGAGTCC 60
 DB 1 CGCGATCGGACCCAAAGCAGGTGCGCGCGCGGCGGAGAGCGCGCGGCGGTGAGTCC 60
 QY 61 TCGACCCCGGTGTCGGGCTAGTCCAGAGCGGAGCGGCGGCGGTGAGGCGGCGGCGG 120
 DB 61 TCGACCCCGGTGTCGGGCTAGTCCAGAGCGGAGCGGCGGCGGTGAGGCGGCGGCGG 120
 QY 121 CCCGCGATGGAGCGGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
 DB 121 CCCGCGATGGAGCGGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
 QY 181 GCGCGCGCGGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
 DB 181 GCGCGCGCGGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
 QY 241 ACTGTGGGACATCGAGAGCTGCTGCTGATGATGAGGTGAGGCTACCCCGGAGCT 300

DB 241 ACTGTGGGCAACATCGAGAGCTGGCTGCTGATGTAAGAGTGCAGGCTACCCCGGAGCT 300
 QY 301 TTGATCCCTACAGATGTGACCTATCAAAATGAAGAGGACATCTCTTCCATGTTCTCAGCT 360
 DB 301 TTGATCCCTACAGATGTGACCTATCAAAATGAAGAGGACATCTCTTCCATGTTCTCAGCT 360
 QY 361 ATCCGTTCTCAGACAGCGGTGTAGACATCTGCATCAACAATGTGGTGGTGGCGGCGCT 420
 DB 361 ATCCGTTCTCAGACAGCGGTGTAGACATCTGCATCAACAATGTGGTGGTGGCGGCGCT 420
 QY 421 GACACCTGCTCTCAGGAGCAGCAGCTGTTGCGAAGGACATGTTCAATGTGAACGTGCTG 480
 DB 421 GACACCTGCTCTCAGGAGCAGCAGCTGTTGCGAAGGACATGTTCAATGTGAACGTGCTG 480
 QY 481 GGCCTCAGCATCTGCACACCGGAAGCCTACCAAGTCCATGAAGGAGCGGAATGTGGACGAT 540
 DB 481 GGCCTCAGCATCTGCACACCGGAAGCCTACCAAGTCCATGAAGGAGCGGAATGTGGACGAT 540
 QY 541 GGGCACATCAATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGTCTGTGACC 600
 DB 541 GGGCACATCAATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGTCTGTGACC 600
 QY 601 CACTTCTATAGTGCCACCAAGTATGCCGTCACTGCGCTGACAGAGGAGCTGAGGCAAGAG 660
 DB 601 CACTTCTATAGTGCCACCAAGTATGCCGTCACTGCGCTGACAGAGGAGCTGAGGCAAGAG 660
 QY 661 CTTGCGGAGGCCAGACCCACATCCGAGCCAGTGCATCTCTCCAGGTGTGTGGAGACA 720
 DB 661 CTTGCGGAGGCCAGACCCACATCCGAGCCAGTGCATCTCTCCAGGTGTGTGGAGACA 720
 QY 721 CAATTCGGCTTCAAACTCCAGACAGGACCTTGAGAGGAGCTGCCACCTATGAGCA 780
 DB 721 CAATTCGGCTTCAAACTCCAGACAGGACCTTGAGAGGAGCTGCCACCTATGAGCA 780
 QY 781 ATGAAGTGTCTCAAAACCGGAGATGTGCGGAGGTGTATCTAGTCTCTCAGCACCCCC 840
 DB 781 ATGAAGTGTCTCAAAACCGGAGATGTGCGGAGGTGTATCTAGTCTCTCAGCACCCCC 840
 QY 841 GCACATCCAGATTTGGAGACATCCAGATGAGGCCACCGAGAGCTGACCTAGTACTG 900
 DB 841 GCACATCCAGATTTGGAGACATCCAGATGAGGCCACCGAGAGCTGACCTAGTACTG 900
 QY 901 TGGGAGCTCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 960
 DB 901 TGGGAGCT 960
 QY 961 TGTGATTTCTGGATCAGGGATACCACTTCTGTCCACACCCCGAGCGGCTAGAAA 1020
 DB 961 TGTGATTTCTGGATCAGGGATACCACTTCTGTCCACACCCCGAGCGGCTAGAAA 1020
 QY 1021 ATTTGTTGAGATTTTATATCATCTTGTCAAAATGCTTCAAGTGTAAATGTGAAAAATG 1080
 DB 1021 ATTTGTTGAGATTTTATATCATCTTGTCAAAATGCTTCAAGTGTAAATGTGAAAAATG 1080
 QY 1081 GGCTGGGAAAAGAGGTGGTGTCCCTAATCTTTTACTTTTACTTTTACTTTTACTTTTACT 1140
 DB 1081 GGCTGGGAAAAGAGGTGGTGTCCCTAATCTTTTACTTTTACTTTTACTTTTACTTTTACT 1140
 QY 1141 TGGGACCTTGCCCTTTTGTCT 1200
 DB 1141 TGGGACCTTGCCCTTTTGTCT 1200
 QY 1201 GCCAAAATCCCAATCTTCTTTCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1260
 DB 1201 GCCAAAATCCCAATCTTCTTTCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1260
 QY 1261 GAGGCTTTCACCTTATATCTGTGTTGTTATCCAGGCTCCAGACTTCTCTCTCTCTCTCTCT 1320
 DB 1261 GAGGCTTTCACCTTATATCTGTGTTGTTATCCAGGCTCCAGACTTCTCTCTCTCTCTCT 1320
 QY 1321 CCCACTGACCCCTCTCCCGCTTATCTATCTCTTTCTCGGCTCCCGAGCCAGCTCTTGCT 1380
 DB 1321 CCCACTGACCCCTCTCCCGCTTATCTATCTCTTCTCGGCTCCCGAGCCAGCTCTTGCT 1380

QY 1381 TCTGTGTCCTCCTGGGTGATCCCTCCACTCTGACTCTGACTATGGCAGAGAACCA 1440
Db 1381 TCTGTGTCCTCCTGGGTGATCCCTCCACTCTGACTCTGACTATGGCAGAGAACCA 1440
QY 1441 GGGCCTGCCCCAGTGGATTTCATGTGATCATTAAGAAAGAAATCGCAACCAAAAAA 1500
Db 1441 GGGCCTGCCCCAGTGGATTTCATGTGATCATTAAGAAAGAAATCGCAACCAAAAAA 1500
QY 1501 AAAAA 1505
Db 1501 AAAAA 1505
RESULT 2
AAF92121
ID AAF92121 standard; cDNA; 1505 BP.
AC AAF92121;
XX
DT 15-MAY-2001 (first entry)
DE Human PRO1774 cDNA.
XX
KW Human; PRO protein; mapping; ss.
XX
OS Homo sapiens.
XX
PN WO200116318-A2.
PD 08-MAR-2001.
XX
PF 24-AUG-2000; 2000WO-US023328.
XX
PR 01-SEP-1999; 99WO-US020111.
PR 15-SEP-1999; 99WO-US021090.
PR 07-DEC-1999; 99US-0169495P.
PR 09-DEC-1999; 99US-0170262P.
PR 11-JAN-2000; 2000US-0175481P.
PR 18-FEB-2000; 2000WO-US004341.
PR 22-FEB-2000; 2000WO-US004342.
PR 01-MAR-2000; 2000WO-US005601.
PR 03-MAR-2000; 2000US-0187202P.
PR 21-MAR-2000; 2000US-0191007P.
PR 30-MAR-2000; 2000WO-US008439.
PR 25-APR-2000; 2000US-0199397P.
PR 22-MAY-2000; 2000WO-US014042.
PR 05-JUN-2000; 2000US-0209832P.
XX
PA (GETH) GENENTECH INC.
XX
XX Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi CJ, Gurney AL, Watanabe CK, Wood WI;
XX
DR WPI; 2001-183260/18.
DR P-PSDB; AAB87589.
XX
PT Eighty four nucleic acids encoding PRO polypeptides, useful in molecular
PT biology, including use as hybridization probes, and in chromosome and
PT gene mapping.
XX
PS Claim 2; Fig 127; 278pp; English.
XX
CC The present sequence is the coding sequence for a human PRO polypeptide
CC (secreted and transmembrane). The PRO protein, and PRO agonists, PRO
CC antagonists or anti-PRO antibodies are useful for preparation of a
CC medicament useful in the treatment of a condition which is responsive to
CC the PRO protein, agonists, antagonists or anti-PRO antibodies. The PRO
CC protein may also be employed as molecular weight markers for protein
CC electrophoresis. The PRO coding sequence has applications in molecular
CC biology, including use as hybridisation probes, and in chromosome and
CC gene mapping

XX Sequence 1505 BP; 309 A; 442 C; 419 G; 335 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 1505; DB 4; Length 1505;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGCGATCGGACCCAGCAGSTCGCGCGCGCGGCGAGAGAGCGGCCGGGCTGAGCTCC 60
Db 1 CGCGATCGGACCCAGCAGSTCGCGCGCGCGGCGAGAGAGCGGCCGGGCTGAGCTCC 60
QY 61 TCGACCCCGCTGTGCGGCTAGTCCAGCGAGCGGCGGCGGCTGGGCCCATGCGCCAGG 120
Db 61 TCGACCCCGCTGTGCGGCTAGTCCAGCGAGCGGCGGCGGCTGGGCCCATGCGCCAGG 120
QY 121 CCGGCGATGGAGCGGTGGCGCGGCGGCTGCGTGGCTGAGCGGGGGCTCGGGGGGCATC 180
Db 121 CCGGCGATGGAGCGGTGGCGCGGCGGCTGCGTGGCTGAGCGGGGGCTCGGGGGGCATC 180
QY 181 GCGCGGCGCTGGCGCGGCGGCGGCTGGTCCAGCAGGACTGAAGGTGGTGGGCTGGCGCGC 240
Db 181 GCGCGGCGCTGGCGCGGCGGCGGCTGGTCCAGCAGGACTGAAGGTGGTGGGCTGGCGCGC 240
QY 241 ACTGTGGCAACATCGAGGAGCTGGCTGCTGAATGTAAGAGTGCAGGCTACCCCGGACT 300
Db 241 ACTGTGGCAACATCGAGGAGCTGGCTGCTGAATGTAAGAGTGCAGGCTACCCCGGACT 300
QY 301 TTGATCCCGTACAGATGACCTATCAAAATGAAGGACATCTCTTCCATGTTCTCAGCT 360
Db 301 TTGATCCCGTACAGATGACCTATCAAAATGAAGGACATCTCTTCCATGTTCTCAGCT 360
QY 361 ATCGGTTCTGACGACGCGGTGTAGACATCTGCATCAACAATGCTGGCTGGCGCGGCT 420
Db 361 ATCGGTTCTGACGACGCGGTGTAGACATCTGCATCAACAATGCTGGCTGGCGCGGCT 420
QY 421 GACACCTGCTCTCAGGCGAGCACAGTGGTGGAGGACATGTTCAATGTCAACGTGCTG 480
Db 421 GACACCTGCTCTCAGGCGAGCACAGTGGTGGAGGACATGTTCAATGTCAACGTGCTG 480
QY 481 GCGCTCAGCATCTGCACACGCGGAGCCTACCACTCCATGAAGGAGCGGAAATGTGACAT 540
Db 481 GCGCTCAGCATCTGCACACGCGGAGCCTACCACTCCATGAAGGAGCGGAAATGTGACAT 540
QY 541 GGGCAGCATTAATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGCTGTGACC 600
Db 541 GGGCAGCATTAATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGCTGTGACC 600
QY 601 CACTTCTATAGTGCCACCAAGTATGCCGTCACTGGCTGACAGAGGAGCTGAGGCAAGAG 660
Db 601 CACTTCTATAGTGCCACCAAGTATGCCGTCACTGGCTGACAGAGGAGCTGAGGCAAGAG 660
QY 661 CTTGGGAGGCCAGACCCACATCCGAGCCAGTGCATCTCTCCAGGTGTGTGGAGACA 720
Db 661 CTTGGGAGGCCAGACCCACATCCGAGCCAGTGCATCTCTCCAGGTGTGTGGAGACA 720
QY 721 CAATTGCGCTTCAAACTCCAGCAGACCCCTGAGAGGCGAGCTGCCACCTATCAGCAA 780
Db 721 CAATTGCGCTTCAAACTCCAGCAGACCCCTGAGAGGCGAGCTGCCACCTATCAGCAA 780
QY 781 ATGAAGTGTCTCAAAACCGAGGATGTGGCCGAGGCTGTTATCTACGTCTCAGCACCCCC 840
Db 781 ATGAAGTGTCTCAAAACCGAGGATGTGGCCGAGGCTGTTATCTACGTCTCAGCACCCCC 840
QY 841 GCACATCCAGATTGGAGACATCCAGATGAGGCCACCGAGAGAGTGCACCTAGTGAATG 900
Db 841 GCACATCCAGATTGGAGACATCCAGATGAGGCCACCGAGAGAGTGCACCTAGTGAATG 900
QY 901 TGGGAGCTCTCTTCCCTCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db 901 TGGGAGCTCTCTTCCCTCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
QY 961 TGTTGATTTCTTGATTCAGGGATACCACTTCTCTGTCACACCCCGAGCGGGCTAGAAA 1020
Db 961 TGTTGATTTCTTGATTCAGGGATACCACTTCTCTGTCACACCCCGAGCGGGCTAGAAA 1020

QY	1381	TCTTGTCCCTCTCTGGGTCA	TCCCTCCACTCTGACTCTGACTATGGCAGCAGAACCA	1440	PR	21-APR-1998;	98US-0082569P.
					PR	22-APR-1998;	98US-0082704P.
Db	1381	TCTTGTCCCTCTCTGGGTCA	TCCCTCCACTCTGACTCTGACTATGGCAGCAGAACCA	1440	PR	22-APR-1998;	98US-0082797P.
					PR	28-APR-1998;	98US-0083322P.
QY	1441	GGGCTGGCCAGTGGATTTCAT	TGGTGATCATTAAGAAAAATCGCAACCAAAAAA	1500	PR	29-APR-1998;	98US-0083495P.
					PR	29-APR-1998;	98US-0083496P.
Db	1441	GGGCTGGCCAGTGGATTTCAT	TGGTGATCATTAAGAAAAATCGCAACCAAAAAA	1500	PR	29-APR-1998;	98US-0083499P.
					PR	29-APR-1998;	98US-0083559P.
QY	1501	AAAAA	1505	PR	05-MAY-1998;	98US-0084366P.	98US-0084366P.
				PR	06-MAY-1998;	98US-0084414P.	98US-0084414P.
Db	1501	AAAAA	1505	PR	07-MAY-1998;	98US-0084639P.	98US-0084639P.
				PR	07-MAY-1998;	98US-0084640P.	98US-0084640P.
				PR	07-MAY-1998;	98US-0084643P.	98US-0084643P.
				PR	15-MAY-1998;	98US-0085579P.	98US-0085579P.
				PR	15-MAY-1998;	98US-0085580P.	98US-0085580P.
				PR	15-MAY-1998;	98US-0085582P.	98US-0085582P.
				PR	15-MAY-1998;	98US-0085700P.	98US-0085700P.
				PR	18-MAY-1998;	98US-0086023P.	98US-0086023P.
				PR	22-MAY-1998;	98US-0086392P.	98US-0086392P.
				PR	22-MAY-1998;	98US-0086486P.	98US-0086486P.
				PR	28-MAY-1998;	98US-0087098P.	98US-0087098P.
				PR	28-MAY-1998;	98US-0087208P.	98US-0087208P.
				PR	02-JUN-1998;	98US-0087609P.	98US-0087609P.
				PR	03-JUN-1998;	98US-0087759P.	98US-0087759P.
				PR	03-JUN-1998;	98US-0087827P.	98US-0087827P.
				PR	04-JUN-1998;	98US-0088025P.	98US-0088025P.
				PR	04-JUN-1998;	98US-0088028P.	98US-0088028P.
				PR	04-JUN-1998;	98US-0088033P.	98US-0088033P.
				PR	04-JUN-1998;	98US-0088167P.	98US-0088167P.
				PR	05-JUN-1998;	98US-0088202P.	98US-0088202P.
				PR	05-JUN-1998;	98US-0088212P.	98US-0088212P.
				PR	05-JUN-1998;	98US-0088655P.	98US-0088655P.
				PR	09-JUN-1998;	98US-0088722P.	98US-0088722P.
				PR	10-JUN-1998;	98US-0088738P.	98US-0088738P.
				PR	10-JUN-1998;	98US-0088740P.	98US-0088740P.
				PR	10-JUN-1998;	98US-0088811P.	98US-0088811P.
				PR	10-JUN-1998;	98US-0088824P.	98US-0088824P.
				PR	10-JUN-1998;	98US-0088825P.	98US-0088825P.
				PR	11-JUN-1998;	98US-0088861P.	98US-0088861P.
				PR	11-JUN-1998;	98US-0088863P.	98US-0088863P.
				PR	11-JUN-1998;	98US-0088876P.	98US-0088876P.
				PR	12-JUN-1998;	98US-0089090P.	98US-0089090P.
				PR	12-JUN-1998;	98US-0089105P.	98US-0089105P.
				PR	16-JUN-1998;	98US-0089512P.	98US-0089512P.
				PR	16-JUN-1998;	98US-0089514P.	98US-0089514P.
				PR	17-JUN-1998;	98US-0089538P.	98US-0089538P.
				PR	17-JUN-1998;	98US-0089598P.	98US-0089598P.
				PR	17-JUN-1998;	98US-0089653P.	98US-0089653P.
				PR	18-JUN-1998;	98US-0089908P.	98US-0089908P.
				PR	19-JUN-1998;	98US-0089952P.	98US-0089952P.
				PR	22-JUN-1998;	98US-0090246P.	98US-0090246P.
				PR	22-JUN-1998;	98US-0090252P.	98US-0090252P.
				PR	22-JUN-1998;	98US-0090254P.	98US-0090254P.
				PR	24-JUN-1998;	98US-0090429P.	98US-0090429P.
				PR	24-JUN-1998;	98US-0090435P.	98US-0090435P.
				PR	24-JUN-1998;	98US-0090444P.	98US-0090444P.
				PR	24-JUN-1998;	98US-0090461P.	98US-0090461P.
				PR	24-JUN-1998;	98US-0090535P.	98US-0090535P.
				PR	24-JUN-1998;	98US-0090540P.	98US-0090540P.
				PR	25-JUN-1998;	98US-0090678P.	98US-0090678P.
				PR	25-JUN-1998;	98US-0090688P.	98US-0090688P.
				PR	25-JUN-1998;	98US-0090690P.	98US-0090690P.
				PR	25-JUN-1998;	98US-0090694P.	98US-0090694P.
				PR	25-JUN-1998;	98US-0090695P.	98US-0090695P.
				PR	25-JUN-1998;	98US-0090696P.	98US-0090696P.
				PR	26-JUN-1998;	98US-00105413.	98US-00105413.
				PR	26-JUN-1998;	98US-0090862P.	98US-0090862P.
				PR	26-JUN-1998;	98US-0090863P.	98US-0090863P.

[illegible]

QY	1081	GGCTGGGGAAGAGGTGGTGCCCTAAATGTTTACTTGTAACTGTGTTCTGTGCCCC	1140	PR	24-NOV-1997;	97US-0066466P.
Db	1081	GGCTGGGGAAGAGGTGGTGCCCTAAATGTTTACTTGTAACTGTGTTCTGTGCCCC	1140	PR	24-NOV-1997;	97US-0066772P.
QY	1141	TGGGCACCTTGGCCTTTGTCTGTCTCTCAGTGTCTTCCCTTTGACATGGGAAAGAGTTGTG	1200	PR	11-DEC-1997;	97US-0069333P.
Db	1141	TGGGCACCTTGGCCTTTGTCTGTCTCTCAGTGTCTTCCCTTTGACATGGGAAAGAGTTGTG	1200	PR	12-DEC-1997;	97US-0069425P.
QY	1201	GCAGAAATCCCATCTTCTTGGACCTCAACGCTGTGGCTCAGGGCTGGGGTGGCAGAGG	1260	PR	17-DEC-1997;	97US-0069870P.
Db	1201	GCAGAAATCCCATCTTCTTGGACCTCAACGCTGTGGCTCAGGGCTGGGGTGGCAGAGG	1260	PR	18-DEC-1997;	97US-0068017P.
QY	1261	GAGGCTTCACTTATATCTGTGTGTATCCAGGCTCCAGACTTCTCCTCTGCTCGC	1320	PR	10-MAR-1998;	98US-0077450P.
Db	1261	GAGGCTTCACTTATATCTGTGTGTATCCAGGCTCCAGACTTCTCCTCTGCTCGC	1320	PR	11-MAR-1998;	98US-0077632P.
QY	1321	CCCACTGACCTCTCCCCCTTATCTATCTCTCGGCTCCCAAGCCAGTCTTGGCT	1380	PR	11-MAR-1998;	98US-0077649P.
Db	1321	CCCACTGACCTCTCCCCCTTATCTATCTCTCGGCTCCCAAGCCAGTCTTGGCT	1380	PR	20-MAR-1998;	98US-0078886P.
QY	1381	TCTTGTCCCTCTCGTGGGTCACTCCCTCACTCTGACTCTGACTATGGCAGCAGACCA	1440	PR	20-MAR-1998;	98US-0078939P.
Db	1381	TCTTGTCCCTCTCGTGGGTCACTCCCTCACTCTGACTCTGACTATGGCAGCAGACCA	1440	PR	27-MAR-1998;	98US-0079664P.
QY	1441	GGGCTGGCCAGTGGATTTCATGGTGATCAATTAAGAAAGAAATCGCAACCAAAAA	1500	PR	31-MAR-1998;	98US-0080107P.
Db	1441	GGGCTGGCCAGTGGATTTCATGGTGATCAATTAAGAAAGAAATCGCAACCAAAAA	1500	PR	01-APR-1998;	98US-0080194P.
QY	1501	AAAAA 1505		PR	01-APR-1998;	98US-0080327P.
Db	1501	AAAAA 1505		PR	01-APR-1998;	98US-0080333P.
RESULT 5				PR	08-APR-1998;	98US-0081049P.
ACA75688				PR	08-APR-1998;	98US-0081070P.
ID ACA75688 standard; cdNA; 1505 BP.				PR	09-APR-1998;	98US-0081195P.
XX				PR	15-APR-1998;	98US-0081838P.
ACA75688;				PR	21-APR-1998;	98US-0082568P.
DT 07-JUL-2003 (first entry)				PR	21-APR-1998;	98US-0082569P.
XX				PR	22-APR-1998;	98US-0082704P.
DE				PR	22-APR-1998;	98US-0082797P.
KW				PR	28-APR-1998;	98US-0083322P.
KW				PR	29-APR-1998;	98US-0083495P.
KW				PR	29-APR-1998;	98US-0083496P.
KW				PR	29-APR-1998;	98US-0083499P.
KW				PR	29-APR-1998;	98US-0083559P.
XX				PR	05-MAY-1998;	98US-0084366P.
OS				PR	06-MAY-1998;	98US-0084414P.
XX				PR	07-MAY-1998;	98US-0084639P.
PN				PR	07-MAY-1998;	98US-0084640P.
PD				PR	07-MAY-1998;	98US-0084643P.
XX				PR	15-MAY-1998;	98US-0085579P.
XX				PR	15-MAY-1998;	98US-0085580P.
XX				PR	15-MAY-1998;	98US-0085582P.
XX				PR	15-MAY-1998;	98US-0085700P.
XX				PR	18-MAY-1998;	98US-0086023P.
XX				PR	22-MAY-1998;	98US-0086392P.
XX				PR	22-MAY-1998;	98US-0086486P.
XX				PR	28-MAY-1998;	98US-0087098P.
XX				PR	28-MAY-1998;	98US-0087208P.
XX				PR	02-JUN-1998;	98US-0087609P.
XX				PR	02-JUN-1998;	98US-0087759P.
XX				PR	03-JUN-1998;	98US-0087827P.
XX				PR	04-JUN-1998;	98US-0088025P.
XX				PR	04-JUN-1998;	98US-0088028P.
XX				PR	04-JUN-1998;	98US-0088029P.
XX				PR	04-JUN-1998;	98US-0088031P.
XX				PR	04-JUN-1998;	98US-0088328P.
XX				PR	05-JUN-1998;	98US-0088167P.
XX				PR	05-JUN-1998;	98US-0088202P.
XX				PR	05-JUN-1998;	98US-0088212P.
XX				PR	05-JUN-1998;	98US-0088217P.
XX				PR	09-JUN-1998;	98US-0088655P.
XX				PR	10-JUN-1998;	98US-0088722P.
XX				PR	10-JUN-1998;	98US-0088738P.
XX				PR	10-JUN-1998;	98US-0088740P.
XX				PR	10-JUN-1998;	98US-0088811P.
XX				PR	10-JUN-1998;	98US-0088824P.
XX				PR	10-JUN-1998;	98US-0088825P.
XX				PR	10-JUN-1998;	98US-0088826P.
XX				PR	11-JUN-1998;	98US-0088861P.
XX				PR	11-JUN-1998;	98US-0088863P.
XX				PR	11-JUN-1998;	98US-0088876P.
XX				PR	12-JUN-1998;	98US-0089090P.
XX				PR	12-JUN-1998;	98US-0089105P.
XX				PR	16-JUN-1998;	98US-0089512P.
XX				PR	16-JUN-1998;	98US-0089514P.
XX				PR	17-JUN-1998;	98US-0089538P.
XX				PR	17-JUN-1998;	98US-0089598P.

Novel human secreted and transmembrane protein PRO1774 cdNA.
Human; secreted and transmembrane protein; PRO; Gene therapy;
tumour necrosis factor-alpha release; TNF-alpha release;
chondrocyte proliferation; chondrocyte differentiation; tumour;
adrenal tumour; lung tumour; colon tumour; breast tumour;
prostate tumour; rectal tumour; cervical tumour; liver tumour; gene; ss.

[illegible]

Qy	481	GCCTCTAGCATCTCGACACGCGGAAGCCCTACCACTGTCCTATGAAGGAGCGGAATGTGGACGAT	540
Db	481	GCCTCTAGCATCTCGACACGCGGAAGCCCTACCACTGTCCTATGAAGGAGCGGAATGTGGACGAT	540
Qy	541	GGGCACATCATTAACATCAATAGCATCTCTGGCCACCGAGTGTGTACCCCTGTCTGTGACC	600
Db	541	GGGCACATCATTAACATCAATAGCATCTCTGGCCACCGAGTGTGTACCCCTGTCTGTGACC	600
Qy	601	CACTTCTATAGTGGCCACCAAGTATGCGGTCACTGCGCTGACAGAGGACTGAGCGAAGAG	660
Db	601	CACTTCTATAGTGGCCACCAAGTATGCGGTCACTGCGCTGACAGAGGACTGAGCGAAGAG	660
Qy	661	CTTCGGGAGGCCAGACCCACATCCGAGCCACGTGCACTCTCCAGGTGTGGTGAGACA	720
Db	661	CTTCGGGAGGCCAGACCCACATCCGAGCCACGTGCACTCTCCAGGTGTGGTGAGACA	720
Qy	721	CAATTCGGCTTCAAACTCCAGCAAGAGCCCTCAGAAAGGAGCTGCCACTATGAGCAA	780
Db	721	CAATTCGGCTTCAAACTCCAGCAAGAGCCCTCAGAAAGGAGCTGCCACTATGAGCAA	780
Qy	781	ATGAAGTGTCTCAAAACCGAGGATGTGGCCGAGGCTGTATCTACGTCTTCAGCACCCCC	840
Db	781	ATGAAGTGTCTCAAAACCGAGGATGTGGCCGAGGCTGTATCTACGTCTTCAGCACCCCC	840
Qy	841	GCACACATCCAGATTTGAGACATCCAGATGAGGCCCCAGGAGCAGGTGACCTAGTGA	900
Db	841	GCACACATCCAGATTTGAGACATCCAGATGAGGCCCCAGGAGCAGGTGACCTAGTGA	900
Qy	901	TGGGAGCTCCCTCTCCCTCCCCACCCCTCATATGCTTGCCTCTCCCTCTGGATTTT	960
Db	901	TGGGAGCTCCCTCTCCCTCCCCACCCCTCATATGCTTGCCTCTCCCTCTGGATTTT	960
Qy	961	TGTTGATTTCTGATCAAGGATACCACTTCCTCTGCCACCCCGACCGAGGGCTAGAAA	1020
Db	961	TGTTGATTTCTGATCAAGGATACCACTTCCTCTGCCACCCCGACCGAGGGCTAGAAA	1020
Qy	1021	ATTTGTTTGAGATTTTATATCATCTGTGCAAAATGCTTCAGTTGTAAATGTGAAAAATG	1080
Db	1021	ATTTGTTTGAGATTTTATATCATCTGTGCAAAATGCTTCAGTTGTAAATGTGAAAAATG	1080
Qy	1081	GCCTGGGAAAGGAGGTGGTCTCCCTAACTGTTTACTTGTAACTTGTCTCTGTGCCCC	1140
Db	1081	GCCTGGGAAAGGAGGTGGTCTCCCTAACTGTTTACTTGTAACTTGTCTCTGTGCCCC	1140
Qy	1141	TGGGCATCTGGCCCTTTGTCCTCTCAGTGTCTTCCCTTTGACATGGGAAGGAGTTGTG	1200
Db	1141	TGGGCATCTGGCCCTTTGTCCTCTCAGTGTCTTCCCTTTGACATGGGAAGGAGTTGTG	1200
Qy	1201	GCCAAATCCCACTTCTTGACACTCAACGTCTGTGGCTCAGGGCTGGGGTGGCAGAGG	1260
Db	1201	GCCAAATCCCACTTCTTGACACTCAACGTCTGTGGCTCAGGGCTGGGGTGGCAGAGG	1260
Qy	1261	GAGGCCTTCACCTTATATCTGTGTTGTTATCCAGGGCTCCAGACTTCTCTCTCTGCTGCTG	1320
Db	1261	GAGGCCTTCACCTTATATCTGTGTTGTTATCCAGGGCTCCAGACTTCTCTCTCTGCTGCTG	1320
Qy	1321	CCCAGTCACCCCTCTCCGCCCTTATCTATCTCCCTCTCGGGCTCCCGAGCCAGTCTTGCT	1380
Db	1321	CCCAGTCACCCCTCTCCGCCCTTATCTATCTCCCTCTCGGGCTCCCGAGCCAGTCTTGCT	1380
Qy	1381	TCCTGTCCCTCTCGGGGTGATCCCTCCACTCTGACTCTGACTGTGCGAGCAGAACCA	1440
Db	1381	TCCTGTCCCTCTCGGGGTGATCCCTCCACTCTGACTCTGACTGTGCGAGCAGAACCA	1440
Qy	1441	GGGCTGGCCCCAGTGGATTTTCATGGTGATCATTAATAAAGAAAAATCGCAACCAAAAAA	1500
Db	1441	GGGCTGGCCCCAGTGGATTTTCATGGTGATCATTAATAAAGAAAAATCGCAACCAAAAAA	1500
Qy	1501	AAAAA 1505	
Db	1501	AAAAA 1505	

RESULT 7	
ACC87696	
ID	ACC87696 standard; cDNA; 1505 BP.
XX	
AC	ACC87696;
XX	
DT	05-AUG-2003 (first entry)
XX	
DE	Human secreted polypeptide PRO1774-encoding cDNA, SEQ ID NO:377
XX	
KW	Human; PRO; secreted protein; transmembrane protein;
KW	extracellular domain; tumour necrosis factor-alpha; TNF-alpha;
KW	chondrocyte; proliferation; differentiation; cartilage disorder
KW	bone disorder; arthritis; sports injury; cancer; tumour; diagnosis
KW	adrenal tumour; lung; colon; breast; prostate; kidney; rectum;
KW	liver; drug screening; transgenic animal; genetic analysis;
KW	anti-arthritic; vulnerary; gene therapy; gene; ss.
XX	
OS	Homo sapiens.
XX	
FN	US2003027278-A1.
XX	
PD	06-FEB-2003.
XX	
PF	21-JUN-2002; 2002US-00176987.
XX	
PR	18-SEP-1997; 97US-0059263P.
PR	18-SEP-1997; 97US-0059266P.
PR	17-OCT-1997; 97US-0062250P.
PR	21-OCT-1997; 97US-0063486P.
PR	24-OCT-1997; 97US-0063120P.
PR	24-OCT-1997; 97US-0063121P.
PR	28-OCT-1997; 97US-0063540P.
PR	28-OCT-1997; 97US-0063541P.
PR	28-OCT-1997; 97US-0063544P.
PR	28-OCT-1997; 97US-0063564P.
PR	29-OCT-1997; 97US-0063734P.
PR	31-OCT-1997; 97US-0063870P.
PR	31-OCT-1997; 97US-0064103P.
PR	13-NOV-1997; 97US-0065311P.
PR	21-NOV-1997; 97US-0066120P.
PR	24-NOV-1997; 97US-0066466P.
PR	24-NOV-1997; 97US-0066772P.
PR	11-DEC-1997; 97US-0069335P.
PR	12-DEC-1997; 97US-0069425P.
PR	17-DEC-1997; 97US-0069870P.
PR	18-DEC-1997; 97US-0068017P.
PR	10-MAR-1998; 98US-0077450P.
PR	11-MAR-1998; 98US-0077632P.
PR	11-MAR-1998; 98US-0077649P.
PR	20-MAR-1998; 98US-0078886P.
PR	20-MAR-1998; 98US-0078939P.
PR	27-MAR-1998; 98US-0079664P.
PR	27-MAR-1998; 98US-0079786P.
PR	31-MAR-1998; 98US-0080107P.
PR	31-MAR-1998; 98US-0080194P.
PR	01-APR-1998; 98US-0080327P.
PR	01-APR-1998; 98US-0080333P.
PR	08-APR-1998; 98US-0081049P.
PR	08-APR-1998; 98US-0081070P.
PR	09-APR-1998; 98US-0081195P.
PR	15-APR-1998; 98US-0081838P.
PR	21-APR-1998; 98US-0082568P.
PR	21-APR-1998; 98US-0082569P.
PR	22-APR-1998; 98US-0082704P.
PR	22-APR-1998; 98US-0082797P.
PR	28-APR-1998; 98US-0083322P.
PR	29-APR-1998; 98US-0083495P.
PR	29-APR-1998; 98US-0083496P.
PR	29-APR-1998; 98US-0083499P.
PR	29-APR-1998; 98US-0083559P.
PR	05-MAY-1998; 98US-0084366P.

QY 181 GCGCGGCGGTGCGCGGCGCTGTCCAGCAGGAGCTGAAGGTGGTGGGCTGCGCGCGC 240
Db 181 GCGCGGCGGTGCGCGGCGCTGTCCAGCAGGAGCTGAAGGTGGTGGGCTGCGCGCGC 240
QY 241 ACTGTGGGCAACATCAGAGAGCTGGCTGCTGAATGTAAGAGTGCAGGCTACCCCGGGACT 300
Db 241 ACTGTGGGCAACATCAGAGAGCTGGCTGCTGAATGTAAGAGTGCAGGCTACCCCGGGACT 300
QY 301 TTGATCCCTCAGATGTGACCTATCAATGAAGAGGACATCCTCTCCATGTTCTCAGCT 360
Db 301 TTGATCCCTCAGATGTGACCTATCAATGAAGAGGACATCCTCTCCATGTTCTCAGCT 360
QY 361 ATCCGTTCTCAGCAGCGGCTGACATCTGCATCAACAATGCTTGGCTTGGCCCGGCT 420
Db 361 ATCCGTTCTCAGCAGCGGCTGACATCTGCATCAACAATGCTTGGCTTGGCCCGGCT 420
QY 421 GACACCTCTGCTCAGGAGCAGCACAGTGGTTCGAAGGACATGTTCAATGTGAACGTGCTG 480
Db 421 GACACCTCTGCTCAGGAGCAGCACAGTGGTTCGAAGGACATGTTCAATGTGAACGTGCTG 480
QY 481 GCCCTCAGATCTGCACAGGGAAGCCTACCACTCATGAAGAGCGGAATGTGACGAT 540
Db 481 GCCCTCAGATCTGCACAGGGAAGCCTACCACTCATGAAGAGCGGAATGTGACGAT 540
QY 541 GGGCACATCAATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGTCTGTGACC 600
Db 541 GGGCACATCAATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGTCTGTGACC 600
QY 601 CACTTCTTAGTGCCACCAAGTATGCCGTCACTGCGTGCACAGAGGAGCTGAGCAAGAG 660
Db 601 CACTTCTTAGTGCCACCAAGTATGCCGTCACTGCGTGCACAGAGGAGCTGAGCAAGAG 660
QY 661 CTTCCGGAGGCCAGACCCACATCCGAGCCACGTGCATCTCTCCAGGTGTGTGGAGACA 720
Db 661 CTTCCGGAGGCCAGACCCACATCCGAGCCACGTGCATCTCTCCAGGTGTGTGGAGACA 720
QY 721 CAATTCGGCTTCAAACTCCACACAGGACCTGAGAGGCGAGCTGCCACCTATGAGCAA 780
Db 721 CAATTCGGCTTCAAACTCCACACAGGACCTGAGAGGCGAGCTGCCACCTATGAGCAA 780
QY 781 ATGAAGTGTCTCAAAACCCGAGGATGTGGCCGAGGCTGTTATCTACGTCTCAGCACCCCC 840
Db 781 ATGAAGTGTCTCAAAACCCGAGGATGTGGCCGAGGCTGTTATCTACGTCTCAGCACCCCC 840
QY 841 GCACACATCCAGATTGGAGACATCCAGATGAGGCCCCACGAGCGAGTGCACCTAGTGACTG 900
Db 841 GCACACATCCAGATTGGAGACATCCAGATGAGGCCCCACGAGCGAGTGCACCTAGTGACTG 900
QY 901 TGGGAGCTCCTCTTCCCTCCACCCCTTCATGGCTTGCTCCTGCTCTGATTTTAGG 960
Db 901 TGGGAGCTCCTCTTCCCTCCACCCCTTCATGGCTTGCTCCTGCTCTGATTTTAGG 960
QY 961 TGTGATTTCTGGATCAGGGATACCACTTCTGTCCACACCCCGAGCGGCTAGAAA 1020
Db 961 TGTGATTTCTGGATCAGGGATACCACTTCTGTCCACACCCCGAGCGGCTAGAAA 1020
QY 1021 ATTTGTTGAGATTTTATATCATCTGTGCAAAATGCTTCAGTTGTAATGTGAAAATG 1080
Db 1021 ATTTGTTGAGATTTTATATCATCTGTGCAAAATGCTTCAGTTGTAATGTGAAAATG 1080
QY 1081 GGCTGGGAAAGAGGTGGTCCCTAAATGTTTACTGTTAACTGTTCTGTGCCCC 1140
Db 1081 GGCTGGGAAAGAGGTGGTGGTCCCTAAATGTTTACTGTTAACTGTTCTGTGCCCC 1140
QY 1141 TGGGCACTTGGCCTTGTCTGCTCTCAGTGTCTTCCCTTTCAGATGGGAAAGAGTTGTG 1200
Db 1141 TGGGCACTTGGCCTTGTCTGCTCTCAGTGTCTTCCCTTTCAGATGGGAAAGAGTTGTG 1200
QY 1201 GCCAAATCCCATCTTCTTGACCTCAACGCTGTGGCTCAGGGCTGGGGTGGCAGAG 1260
Db 1201 GCCAAATCCCATCTTCTTGACCTCAACGCTGTGGCTCAGGGCTGGGGTGGCAGAG 1260

QY 1261 GAGGCTTCACTTATATCTGTGTTGTTATCCAGGGCTCCAGACTTCTCTCTGCTGC 1320
Db 1261 GAGGCTTCACTTATATCTGTGTTGTTATCCAGGGCTCCAGACTTCTCTCTGCTGC 1320
QY 1321 CCCACTGCACCTCTCCCCCTTATCTATCTCTTCTCGGCTCCCAGCCCCAGTCTTGGCT 1380
Db 1321 CCCACTGCACCTCTCCCCCTTATCTATCTCTTCTCGGCTCCCAGCCCCAGTCTTGGCT 1380
QY 1381 TCTTGTCCCTCTCGGGGTGATCCCTCAGCTCTGACTCTGACTATGGCAGGAGACCA 1440
Db 1381 TCTTGTCCCTCTCGGGGTGATCCCTCAGCTCTGACTCTGACTATGGCAGGAGACCA 1440
QY 1441 GGGCTGCGCCAGTGGATTTTCATGTGTATCAATTAAGAAAGAAATCGCAACCAAAAA 1500
Db 1441 GGGCTGCGCCAGTGGATTTTCATGTGTATCAATTAAGAAAGAAATCGCAACCAAAAA 1500
QY 1501 AAAAA 1505
Db 1501 AAAAA 1505
RESULT 8
ACCS87082
ID ACC87082 standard; cDNA; 1505 BP.
XX
AC ACC87082;
XX 05-AUG-2003 (first entry)
XX Human secreted polypeptide PRO1774-encoding cDNA, SEQ ID NO:377.
DE Human; PRO; secreted protein; transmembrane protein;
KW extracellular domain; tumour necrosis factor-alpha; TNF-alpha;
KW chondrocyte; proliferation; differentiation; cartilage disorder;
KW bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;
KW adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;
KW liver; drug screening; transgenic animal; genetic analysis;
KW antiarthritic; vulnery; gene therapy; gene; ss.
XX Homo sapiens.
XX US2003036159-A1.
PN 20-FEB-2003.
PD 02-JUL-2002; 2002US-00188773.
PF 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.

PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081193P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083493P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084633P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087603P.
PR 02-JUN-1998; 98US-0087753P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088653P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 12-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 16-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 18-JUN-1998; 98US-0089653P.
PR 19-JUN-1998; 98US-0089908P.
PR 22-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 24-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090423P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.

PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 26-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 18-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096953P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.

PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069333P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0069017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086033P.
PR 22-MAY-1998; 98US-0086322P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 05-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.

PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089588P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 24-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 25-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 26-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-0090697P.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.

XX Homo sapiens.
XX OS
XX XX
PN US2003032113-Al.
XX
PD 13-FEB-2003.
XX PF
XX 20-JUN-2002; 2002US-00176911.
XX
XX 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078866P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081195P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084141P.
PR 06-MAY-1998; 98US-0084366P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086332P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.

PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090439P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 25-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-00977954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.

PR	01-SEP-1998;	98US-0098723P.	QY	361	ATCGTTCTCAGCACAGCGGTGTAGACATCTGCATCAACAATGCTGGCTTGCCCGGGCT	420
PR	02-SEP-1998;	98US-0098803P.	Db	361		
PR	02-SEP-1998;	98US-0098821P.	PR	361	ATCGTTCTCAGCACAGCGGTGTAGACATCTGCATCAACAATGCTGGCTTGCCCGGGCT	420
PR	02-SEP-1998;	98US-0098843P.				
PR	09-SEP-1998;	98US-0099602P.	QY	421	GACACCCCTGCTCTCAGGCAGCACCGAGTGGTGGAGGACATGTTCAATGTGAACCTGCTG	480
PR	10-SEP-1998;	98US-0099741P.	Db	421		
PR	10-SEP-1998;	98US-0099754P.				
PR	10-SEP-1998;	98US-0099763P.	QY	481	GCCTCAGCATCTGCACACCGGAAAGCTTACAGTCCATGAAGGAGCGGAAATGTGACGAT	540
PR	10-SEP-1998;	98US-0099812P.	Db	481	GCCTCAGCATCTGCACACCGGAAAGCTTACAGTCCATGAAGGAGCGGAAATGTGACGAT	540
PR	15-SEP-1998;	98US-0100388P.				
PR	16-SEP-1998;	98US-0100662P.	QY	541	GGGACATCATTAACATCAATAGCATGCTCGCCACCGAGTGTACCCCTGCTCTGACC	600
PR	16-SEP-1998;	98US-0100664P.	Db	541		
PR	16-SEP-1998;	98US-0101751P.	QY	541	GGGACATCATTAACATCAATAGCATGCTCGCCACCGAGTGTACCCCTGCTCTGACC	600
PR	16-SEP-1998;	98WO-US019330.	Db	541	GGGACATCATTAACATCAATAGCATGCTCGCCACCGAGTGTACCCCTGCTCTGACC	600
PR	17-SEP-1998;	98US-0100683P.				
PR	17-SEP-1998;	98US-0100684P.	QY	601	CACCTTCTATAGTGCCACCAAGTATGCCGTCACTGCGCTGACAGAGGGACTGAGGCAAGAG	660
PR	17-SEP-1998;	98US-0100919P.	Db	601	CACCTTCTATAGTGCCACCAAGTATGCCGTCACTGCGCTGACAGAGGGACTGAGGCAAGAG	660
PR	17-SEP-1998;	98US-0100930P.				
PR	18-SEP-1998;	98US-0100849P.	QY	661	CTTCGGAGGCCAGACCCACATCCGAGCCAGTGCATCTCTCCAGGTGTGGTGAGACA	720
PR	18-SEP-1998;	98US-0101014P.	Db	661	CTTCGGAGGCCAGACCCACATCCGAGCCAGTGCATCTCTCCAGGTGTGGTGAGACA	720
PR	18-SEP-1998;	98US-0101068P.				
PR	23-SEP-1998;	98US-0101471P.	QY	721	CAATTTCGCTTCAAACTCCACGACCAAGGACCTGAGAGGAGCTGCCACCTATGAGCAA	780
PR	23-SEP-1998;	98US-0101472P.	Db	721	CAATTTCGCTTCAAACTCCACGACCAAGGACCTGAGAGGAGCTGCCACCTATGAGCAA	780
PR	23-SEP-1998;	98US-0101473P.				
PR	23-SEP-1998;	98US-0101475P.	QY	781	ATGAAGTGTCTCAAAACCCGAGGATGTGCCGAGGCTGTATCTACGTCTCAGCACCCCTC	840
PR	23-SEP-1998;	98US-0101477P.	Db	781	ATGAAGTGTCTCAAAACCCGAGGATGTGCCGAGGCTGTATCTACGTCTCAGCACCCCTC	840
PR	24-SEP-1998;	98US-0101738P.				
PR	24-SEP-1998;	98US-0101739P.	QY	841	GCACATCCAGATTGGAGACATCCAGATGAGGCCACGAGGAGCTGACCTAGTAGCTG	900
PR	24-SEP-1998;	98US-0101739P.	Db	841	GCACATCCAGATTGGAGACATCCAGATGAGGCCACGAGGAGCTGACCTAGTAGCTG	900
PR	24-SEP-1998;	98US-0102331P.				
PR	30-SEP-1998;	98US-0102487P.	QY	901	TGGGAGCTCTCTTCCCTCCACCTTCAATGGCTTGCCTCCTCCCTCTGGATTTTAGG	960
PR	30-SEP-1998;	98US-0102570P.	Db	901	TGGGAGCTCTCTTCCCTCCACCTTCAATGGCTTGCCTCCTCCCTCTGGATTTTAGG	960
PR	30-SEP-1998;	98US-0102571P.				
PR	01-OCT-1998;	98US-0102684P.	QY	961	TGTTGATTCTTGATPCAGGGATACCACTTCTGTCCACACCCGACGAGGGGCTAGAAA	1020
PR	01-OCT-1998;	98US-0102687P.	Db	961	TGTTGATTCTTGATPCAGGGATACCACTTCTGTCCACACCCGACGAGGGGCTAGAAA	1020
PR	02-OCT-1998;	98US-0102965P.				
PR	06-OCT-1998;	98US-0103258P.	QY	1021	ATTTGTTGAGATTTTATATATCTTCTCAAAATGCTTCAGTTTAAATGTGAATAATG	1080
PR	06-OCT-1998;	98US-0103449P.	Db	1021	ATTTGTTGAGATTTTATATATCTTCTCAAAATGCTTCAGTTTAAATGTGAATAATG	1080
PR	07-OCT-1998;	98US-0016897P.				
PR	07-OCT-1998;	98US-0016897P.	QY	1081	GGCTGGGAAAGGAGGTGGTGTCCCTAATTTTACTTTTAACTTTGTCTTGTGCCCC	1140
PR	07-OCT-1998;	98US-0016897P.	Db	1081	GGCTGGGAAAGGAGGTGGTGTCCCTAATTTTACTTTTAACTTTGTCTTGTGCCCC	1140
QY	1	CGCGGATCGGACCCAGCAGGTCCAGCGGCGGCGGCGGAGAGCGCGGGCGCTCAGCTCC				60
Db	1	CGCGGATCGGACCCAGCAGGTCCAGCGGCGGCGGCGGAGAGCGCGGGCGCTCAGCTCC				60
QY	61	TCGACCCCGCTGCGGGCTAGTCCAGCAGGCGGAGCGGCGGCGGCGGCGGCGGCGG				120
Db	61	TCGACCCCGCTGCGGGCTAGTCCAGCAGGCGGAGCGGCGGCGGCGGCGGCGGCGG				120
QY	121	CCCGGATGAGCGGTGGCGGACCGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT				180
Db	121	CCCGGATGAGCGGTGGCGGACCGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT				180
QY	181	GGCGGGCGGCTGGCGGCGGCGGCTGGTCCAGCAGGAGCTGAAGGTGGTGGGCTGGC				240
Db	181	GGCGGGCGGCTGGCGGCGGCGGCTGGTCCAGCAGGAGCTGAAGGTGGTGGGCTGGC				240
QY	241	ACTGTGGGCAACATCCAGAGCTGGCTGCTGAATGTAAAGTGCAGGCTACCCCGGACT				300
Db	241	ACTGTGGGCAACATCCAGAGCTGGCTGCTGAATGTAAAGTGCAGGCTACCCCGGACT				300
QY	301	TTGATCCCTTACAGATGTGACCTTCAATGAAGAGGACATCCTCTCCATGTTCTCAGCT				360
Db	301	TTGATCCCTTACAGATGTGACCTTCAATGAAGAGGACATCCTCTCCATGTTCTCAGCT				360

Db	1441	GGGCTGCCACAGTGGATTTCATGGTGATCATTAAGAAAGAAATCGACACCAAAAA	1500
QY	1501	AAAAA	1505
Db	1501	AAAAA	1505
RESULT 11			
ACA90431			
ID	ACA90431	standard; cDNA; 1505 BP.	
XX	XX		
AC	ACA90431;		
DT	11-AUG-2003	(first entry)	
XX	XX	Novel human secreted and transmembrane protein PRO1774 cDNA.	
DE	XX		
DE	XX	Human; gene therapy; tissue typing; tumour; chondrocyte proliferation;	
KW	XX	chondrocyte differentiation; tumour necrosis factor-alpha release; ss;	
KW	XX	affinity purification; gene.	
XX	XX		
OS	XX	Homo sapiens.	
PN	XX	US2003036147-A1.	
XX	XX		
PD	20-FEB-2003.		
XX	XX		
PF	02-JUL-2002;	2002US-00187741.	
XX	XX		
PR	18-SEP-1997;	97US-0059263P.	
PR	18-SEP-1997;	97US-0059266P.	
PR	17-OCT-1997;	97US-0082250P.	
PR	21-OCT-1997;	97US-0063486P.	
PR	24-OCT-1997;	97US-0063120P.	
PR	24-OCT-1997;	97US-0063121P.	
PR	28-OCT-1997;	97US-0063540P.	
PR	28-OCT-1997;	97US-0063541P.	
PR	28-OCT-1997;	97US-0063544P.	
PR	28-OCT-1997;	97US-0063564P.	
PR	29-OCT-1997;	97US-0063734P.	
PR	31-OCT-1997;	97US-0063870P.	
PR	31-OCT-1997;	97US-0064103P.	
PR	13-NOV-1997;	97US-0065311P.	
PR	21-NOV-1997;	97US-0066120P.	
PR	24-NOV-1997;	97US-0066466P.	
PR	24-NOV-1997;	97US-0066772P.	
PR	11-DEC-1997;	97US-0069335P.	
PR	12-DEC-1997;	97US-0069425P.	
PR	17-DEC-1997;	97US-0069870P.	
PR	18-DEC-1997;	97US-0068017P.	
PR	10-MAR-1998;	98US-0077450P.	
PR	11-MAR-1998;	98US-0077632P.	
PR	11-MAR-1998;	98US-0077649P.	
PR	20-MAR-1998;	98US-0078886P.	
PR	20-MAR-1998;	98US-0078939P.	
PR	27-MAR-1998;	98US-0079664P.	
PR	27-MAR-1998;	98US-0079786P.	
PR	31-MAR-1998;	98US-0080107P.	
PR	31-MAR-1998;	98US-0080194P.	
PR	01-APR-1998;	98US-0080327P.	
PR	01-APR-1998;	98US-0080333P.	
PR	08-APR-1998;	98US-0081049P.	
PR	08-APR-1998;	98US-0081070P.	
PR	09-APR-1998;	98US-0081195P.	
PR	15-APR-1998;	98US-0081838P.	
PR	21-APR-1998;	98US-0082568P.	
PR	21-APR-1998;	98US-0082569P.	
PR	22-APR-1998;	98US-0082704P.	
PR	22-APR-1998;	98US-0082797P.	
PR	28-APR-1998;	98US-0083222P.	
PR	29-APR-1998;	98US-0083495P.	
PR	29-APR-1998;	98US-0083496P.	
PR	29-APR-1998;	98US-0083497P.	
PR	29-APR-1998;	98US-0083498P.	
PR	29-APR-1998;	98US-0083499P.	
PR	29-APR-1998;	98US-0083500P.	
PR	05-MAY-1998;	98US-0084366P.	
PR	06-MAY-1998;	98US-0084414P.	
PR	07-MAY-1998;	98US-0084639P.	
PR	07-MAY-1998;	98US-0084640P.	
PR	07-MAY-1998;	98US-0084643P.	
PR	15-MAY-1998;	98US-0085579P.	
PR	15-MAY-1998;	98US-0085580P.	
PR	15-MAY-1998;	98US-0085582P.	
PR	15-MAY-1998;	98US-0085700P.	
PR	18-MAY-1998;	98US-0086023P.	
PR	22-MAY-1998;	98US-0086392P.	
PR	22-MAY-1998;	98US-0086486P.	
PR	28-MAY-1998;	98US-0087098P.	
PR	28-MAY-1998;	98US-0087208P.	
PR	02-JUN-1998;	98US-0087609P.	
PR	02-JUN-1998;	98US-0087759P.	
PR	03-JUN-1998;	98US-0087827P.	
PR	04-JUN-1998;	98US-0088025P.	
PR	04-JUN-1998;	98US-0088028P.	
PR	04-JUN-1998;	98US-0088029P.	
PR	04-JUN-1998;	98US-0088033P.	
PR	04-JUN-1998;	98US-0088326P.	
PR	05-JUN-1998;	98US-0088167P.	
PR	05-JUN-1998;	98US-0088202P.	
PR	05-JUN-1998;	98US-0088212P.	
PR	05-JUN-1998;	98US-0088217P.	
PR	09-JUN-1998;	98US-0088655P.	
PR	10-JUN-1998;	98US-0088722P.	
PR	10-JUN-1998;	98US-0088738P.	
PR	10-JUN-1998;	98US-0088740P.	
PR	10-JUN-1998;	98US-0088811P.	
PR	10-JUN-1998;	98US-0088824P.	
PR	10-JUN-1998;	98US-0088825P.	
PR	11-JUN-1998;	98US-0088826P.	
PR	11-JUN-1998;	98US-0088861P.	
PR	11-JUN-1998;	98US-0088863P.	
PR	11-JUN-1998;	98US-0088876P.	
PR	12-JUN-1998;	98US-0089090P.	
PR	12-JUN-1998;	98US-0089105P.	
PR	16-JUN-1998;	98US-0089512P.	
PR	16-JUN-1998;	98US-0089514P.	
PR	17-JUN-1998;	98US-0089538P.	
PR	17-JUN-1998;	98US-0089589P.	
PR	17-JUN-1998;	98US-0089653P.	
PR	18-JUN-1998;	98US-0089908P.	
PR	19-JUN-1998;	98US-0089952P.	
PR	22-JUN-1998;	98US-0090246P.	
PR	22-JUN-1998;	98US-0090252P.	
PR	22-JUN-1998;	98US-0090254P.	
PR	24-JUN-1998;	98US-0090429P.	
PR	24-JUN-1998;	98US-0090435P.	
PR	24-JUN-1998;	98US-0090444P.	
PR	24-JUN-1998;	98US-0090461P.	
PR	24-JUN-1998;	98US-0090535P.	
PR	24-JUN-1998;	98US-0090540P.	
PR	25-JUN-1998;	98US-0090676P.	
PR	25-JUN-1998;	98US-0090678P.	
PR	25-JUN-1998;	98US-0090688P.	
PR	25-JUN-1998;	98US-0090690P.	
PR	25-JUN-1998;	98US-0090694P.	
PR	25-JUN-1998;	98US-0090695P.	
PR	25-JUN-1998;	98US-0090696P.	
PR	26-JUN-1998;	98US-00105413.	
PR	26-JUN-1998;	98US-0090862P.	
PR	26-JUN-1998;	98US-0090863P.	
PR	26-JUN-1998;	98US-0091010P.	
PR	01-JUL-1998;	98US-0091359P.	
PR	01-JUL-1998;	98US-0091544P.	
PR	02-JUL-1998;	98US-0091478P.	
PR	02-JUL-1998;	98US-0091486P.	
PR	02-JUL-1998;	98US-0091626P.	

PR	02-JUL-1998;	98US-0091628P.	QY	61	TCGACCCCGCTGTCGGGCTAGTCCAGCGAGCGGACGCGCGCTGGCGCCATGCGCCAG	120
PR	02-JUL-1998;	98US-0091632P.	Db	61	TCGACCCCGCTGTCGGGCTAGTCCAGCGAGCGGACGCGCGCTGGCGCCATGCGCCAG	120
PR	04-JUL-1998;	98US-0094006P.	QY	121	CCCGGCATGAGCGGTGCGCGACCGGCTGCGGCTGTGTGACGGGGGCGCTCGGGGGGCATC	180
PR	10-AUG-1998;	98US-0095598P.	Db	121	CCCGGCATGAGCGGTGCGCGACCGGCTGCGGCTGTGTGACGGGGGCGCTCGGGGGGCATC	180
PR	10-AUG-1998;	98US-0096012P.	QY	181	GGCGGCGCGTGGCCCGGGCCCTGGTCCAGCAGGAGCTGAAGGTGGTGGCTGGCCCGGC	240
PR	17-AUG-1998;	98US-0096757P.	Db	181	GGCGGCGCGTGGCCCGGGCCCTGGTCCAGCAGGAGCTGAAGGTGGTGGCTGGCCCGGC	240
PR	17-AUG-1998;	98US-0096766P.	QY	241	ACTGTGGGCAACATCGAGGAGCTGGCTCTGCTGCTGAAATGTAAGTGCAGGCTACCCCGGACT	300
PR	17-AUG-1998;	98US-0096867P.	Db	241	ACTGTGGGCAACATCGAGGAGCTGGCTCTGCTGCTGAAATGTAAGTGCAGGCTACCCCGGACT	300
PR	17-AUG-1998;	98US-0096891P.	QY	301	TTGATCCCTTACAGATGTGACCTATCAAATGAAGAGGACATCCTCTCCATGTTCTCAGCT	360
PR	17-AUG-1998;	98US-0096897P.	Db	301	TTGATCCCTTACAGATGTGACCTATCAAATGAAGAGGACATCCTCTCCATGTTCTCAGCT	360
PR	18-AUG-1998;	98US-0096949P.	QY	361	ATCGGTTCTCAGCAGCGGTGTAGACATCTGCATCAACAATGCTGGCTTGGCCCGGGCT	420
PR	18-AUG-1998;	98US-0096959P.	Db	361	ATCGGTTCTCAGCAGCGGTGTAGACATCTGCATCAACAATGCTGGCTTGGCCCGGGCT	420
PR	18-AUG-1998;	98US-0097022P.	QY	421	GACACCTGCTCTCAGCGAGCACCAGTGGTGGGAGGACATGTTCAATGTGAACGTGCTG	480
PR	18-AUG-1998;	98US-0097022P.	Db	421	GACACCTGCTCTCAGCGAGCACCAGTGGTGGGAGGACATGTTCAATGTGAACGTGCTG	480
PR	26-AUG-1998;	98US-0097952P.	QY	481	CCCTCAGCATCTGCACACGGGAAGCCTACCAAGTCCATGAAGGAGCGGAATGTGGACAT	540
PR	26-AUG-1998;	98US-0097954P.	Db	481	CCCTCAGCATCTGCACACGGGAAGCCTACCAAGTCCATGAAGGAGCGGAATGTGGACAT	540
PR	26-AUG-1998;	98US-0097955P.	QY	541	GGGACATCATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGCTGTGACC	600
PR	26-AUG-1998;	98US-0097971P.	Db	541	GGGACATCATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGCTGTGACC	600
PR	26-AUG-1998;	98US-0097974P.	QY	601	CACCTTATAGTGCACCAAGTATGCGCTACTGCGCTGACAGAGGAGCTGAGGCAAGAG	660
PR	26-AUG-1998;	98US-0098014P.	Db	601	CACCTTATAGTGCACCAAGTATGCGCTACTGCGCTGACAGAGGAGCTGAGGCAAGAG	660
PR	01-SEP-1998;	98US-0098716P.	QY	661	CTTCGGAGGCCAGACCCACATCCGAGCCACGTCGATCTCTCCAGGTGTGGTGAGACA	720
PR	01-SEP-1998;	98US-0098723P.	Db	661	CTTCGGAGGCCAGACCCACATCCGAGCCACGTCGATCTCTCCAGGTGTGGTGAGACA	720
PR	02-SEP-1998;	98US-0098803P.	QY	721	CAATTCGCTTCAAACCTCCAGCAAGACCTCCAGAAAGGAGGCTGCCACCTATGAGCAA	780
PR	02-SEP-1998;	98US-0098821P.	Db	721	CAATTCGCTTCAAACCTCCAGCAAGACCTCCAGAAAGGAGGCTGCCACCTATGAGCAA	780
PR	02-SEP-1998;	98US-0098843P.	QY	781	ATGAAGTGTCTCAAACCGGAGGTGTGGCCGAGGCTGTATCTACGTCTCAGCACCCCC	840
PR	02-SEP-1998;	98US-0098843P.	Db	781	ATGAAGTGTCTCAAACCGGAGGTGTGGCCGAGGCTGTATCTACGTCTCAGCACCCCC	840
PR	09-SEP-1998;	98US-0099602P.	QY	841	GCACATCCAGATTGGAGACATCCAGATGAGGCCCGGAGGAGGTGACCTAGTACTG	900
PR	09-SEP-1998;	98US-0099602P.	Db	841	GCACATCCAGATTGGAGACATCCAGATGAGGCCCGGAGGAGGTGACCTAGTACTG	900
PR	10-SEP-1998;	98US-0099741P.	QY	901	TGGGAGCTCTCCTTCCCTCCCGACCTTCAATGCTGCTCCCTCGCTTGGATTTAGG	960
PR	10-SEP-1998;	98US-0099754P.	Db	901	TGGGAGCTCTCCTTCCCTCCCGACCTTCAATGCTGCTCCCTCGCTTGGATTTAGG	960
PR	10-SEP-1998;	98US-0099763P.	QY	961	TGTTGATTTCTGGATCAGGGATACCACTTCTCTGTCCACACCCCGACAGGGGTAGAAA	1020
PR	10-SEP-1998;	98US-0099812P.	Db	961	TGTTGATTTCTGGATCAGGGATACCACTTCTCTGTCCACACCCCGACAGGGGTAGAAA	1020
PR	16-SEP-1998;	98US-0100662P.	QY	1021	ATTTGTTGAGATTTTATATCATCTTGTCAAAATGCTTCAGTTGTAATGTGAAAAATG	1080
PR	16-SEP-1998;	98US-0100664P.	Db	1021	ATTTGTTGAGATTTTATATCATCTTGTCAAAATGCTTCAGTTGTAATGTGAAAAATG	1080
PR	16-SEP-1998;	98US-0101014P.	QY	1081	GGCTGGGGAAGGAGGTGGTCCCTAATCTTTTACTTGTAACTTGTCTTGTGCCCC	1140
PR	16-SEP-1998;	98US-0101014P.	Db	1081	GGCTGGGGAAGGAGGTGGTCCCTAATCTTTTACTTGTAACTTGTCTTGTGCCCC	1140
PR	16-SEP-1998;	98US-0101068P.	QY	1141	TGGGACCTTGCCCTTTGTCTCTCTCAGTGTCTTCCCTTTTGACATGGGAAAGGAGTGTG	1200
PR	16-SEP-1998;	98US-0101068P.	Db	1141	TGGGACCTTGCCCTTTGTCTCTCTCAGTGTCTTCCCTTTTGACATGGGAAAGGAGTGTG	1200

Query Match 100.0%; Score 1505; DB 7; Length 1505;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGCGATCGGACCAAGCAGGTTCGCGCGCGGCGGAGGAGCGCGCGGCGTCAAGTCC 60
 Db 1 CGCGATCGGACCAAGCAGGTTCGCGCGCGGCGGAGGAGCGCGCGGCGTCAAGTCC 60

Db	1141		TGGGCACTTGGCCCTTTGCTGCTCTCAGTGTCTTCCCTTTGACATGGGAAGAGTTGTG	1200	PR	17-DEC-1997;
					PR	18-DEC-1997;
Qy	1201	GCACAAATCCCATCTCTTTGGCACTCAACGTCCTGGGCTCAGGGCTGGGGTGGCAGAGG	1260	PR	PR	10-MAR-1998;
					PR	11-MAR-1998;
Db	1201	GCACAAATCCCATCTCTTTGGCACTCAACGTCCTGGGCTCAGGGCTGGGGTGGCAGAGG	1260	PR	PR	11-MAR-1998;
					PR	20-MAR-1998;
Qy	1261	GAGGCTTCACCTTATATCTGTGTGTGTTATCCAGGGCTCCAGACTTCCTCTCTGCTGTC	1320	PR	PR	27-MAR-1998;
					PR	27-MAR-1998;
Db	1261	GAGGCTTCACCTTATATCTGTGTGTGTTATCCAGGGCTCCAGACTTCCTCTCTGCTGTC	1320	PR	PR	31-MAR-1998;
					PR	31-MAR-1998;
Qy	1321	CCCACTGCACCTCTCCCCCTTATCTATCTATCTCTTCTGGGCTCCCGACCCAGTCTTGGCT	1380	PR	PR	01-APR-1998;
					PR	01-APR-1998;
Db	1321	CCCACTGCACCTCTCCCCCTTATCTATCTATCTCTTCTGGGCTCCCGACCCAGTCTTGGCT	1380	PR	PR	08-APR-1998;
					PR	08-APR-1998;
Qy	1381	TCCTGTCCCTCTCGGGGTCATCCCTCCACTCTGACTCTGACTATGGCAGCAGAACACCA	1440	PR	PR	09-APR-1998;
					PR	15-APR-1998;
Db	1381	TCCTGTCCCTCTCGGGGTCATCCCTCCACTCTGACTCTGACTATGGCAGCAGAACACCA	1440	PR	PR	21-APR-1998;
					PR	21-APR-1998;
Qy	1441	GGGCTTGGCCAGTGGATTTTCATGGTATCATTTAAAAAGAAAAATCGCAACCAAAAAA	1500	PR	PR	22-APR-1998;
					PR	22-APR-1998;
Db	1441	GGGCTTGGCCAGTGGATTTTCATGGTATCATTTAAAAAGAAAAATCGCAACCAAAAAA	1500	PR	PR	28-APR-1998;
					PR	28-APR-1998;
Qy	1501	AAAAA 1505		PR	PR	29-APR-1998;
					PR	29-APR-1998;
Db	1501	AAAAA 1505		PR	PR	29-APR-1998;
					PR	29-APR-1998;
RESULT 12						
ACC89538						
ID	ACC89538	standard; cDNA; 1505 BP.				
XX	ACC89538;					
XX	ACC89538;					
DT	11-AUG-2003	(first entry)				
XX						
DE	Human secreted polypeptide PRO1774-encoding cDNA, SEQ ID NO:377.					
XX						
KW	Human; PRO; secreted protein; transmembrane protein;					
KW	extracellular domain; tumour necrosis factor-alpha; TNF-alpha;					
KW	chondrocyte; proliferation; differentiation; cartilage disorder;					
KW	bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;					
KW	adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;					
KW	liver; drug screening; transgenic animal; genetic analysis;					
KW	antiarthritic; vulneryary; gene therapy; gene; ss.					
OS	Homo sapiens.					
PF						
FN	US2003027264-A1.					
PD	06-FEB-2003.					
XX						
XX	18-JUN-2002; 2002US-00174579.					
PF						
XX						
PR	18-SEP-1997; 97US-0059263P.					
PR	18-SEP-1997; 97US-0059266P.					
PR	17-OCT-1997; 97US-0062250P.					
PR	21-OCT-1997; 97US-0063486P.					
PR	24-OCT-1997; 97US-0063120P.					
PR	24-OCT-1997; 97US-0063121P.					
PR	28-OCT-1997; 97US-0063540P.					
PR	28-OCT-1997; 97US-0063541P.					
PR	28-OCT-1997; 97US-0063544P.					
PR	28-OCT-1997; 97US-0063564P.					
PR	29-OCT-1997; 97US-0063734P.					
PR	31-OCT-1997; 97US-0063870P.					
PR	31-OCT-1997; 97US-0064103P.					
PR	13-NOV-1997; 97US-0065311P.					
PR	21-NOV-1997; 97US-0066120P.					
PR	24-NOV-1997; 97US-0066466P.					
PR	24-NOV-1997; 97US-0066772P.					
PR	11-DEC-1997; 97US-0069335P.					
PR	12-DEC-1997; 97US-0069425P.					
97US-0069870P;						
97US-0068017P;						
98US-0077450P;						
98US-0077632P;						
98US-						

```
PR 22-JUN-1998; 98US-0090252P;
PR 22-JUN-1998; 98US-0090254P;
PR 24-JUN-1998; 98US-0090429P;
PR 24-JUN-1998; 98US-0090435P;
PR 24-JUN-1998; 98US-0090441P;
PR 24-JUN-1998; 98US-0090446P;
PR 24-JUN-1998; 98US-0090453P;
PR 24-JUN-1998; 98US-0090540P;
PR 25-JUN-1998; 98US-0090546P;
PR 25-JUN-1998; 98US-0090678P;
PR 25-JUN-1998; 98US-0090688P;
PR 25-JUN-1998; 98US-0090690P;
PR 25-JUN-1998; 98US-0090694P;
PR 25-JUN-1998; 98US-0090695P;
PR 25-JUN-1998; 98US-0090696P;
PR 26-JUN-1998; 98US-00105413;
PR 26-JUN-1998; 98US-0090862P;
PR 26-JUN-1998; 98US-0090863P;
PR 26-JUN-1998; 98US-0091010P;
PR 01-JUL-1998; 98US-0091359P;
PR 01-JUL-1998; 98US-0091544P;
PR 02-JUL-1998; 98US-0091478P;
PR 02-JUL-1998; 98US-0091486P;
PR 02-JUL-1998; 98US-0091626P;
PR 02-JUL-1998; 98US-0091628P;
PR 02-JUL-1998; 98US-0091632P;
PR 24-JUL-1998; 98US-0094006P;
PR 04-AUG-1998; 98US-0095282P;
PR 10-AUG-1998; 98US-0095998P;
PR 10-AUG-1998; 98US-0096012P;
PR 17-AUG-1998; 98US-0096757P;
PR 17-AUG-1998; 98US-0096766P;
PR 17-AUG-1998; 98US-0096867P;
PR 17-AUG-1998; 98US-0096891P;
PR 17-AUG-1998; 98US-0096897P;
PR 18-AUG-1998; 98US-0096949P;
PR 18-AUG-1998; 98US-0096959P;
PR 26-AUG-1998; 98US-0097022P;
PR 26-AUG-1998; 98US-0097952P;
PR 26-AUG-1998; 98US-0097954P;
PR 26-AUG-1998; 98US-0097955P;
PR 26-AUG-1998; 98US-0097971P;
PR 26-AUG-1998; 98US-0097974P;
PR 26-AUG-1998; 98US-0098014P;
PR 01-SEP-1998; 98US-0098716P;
PR 01-SEP-1998; 98US-0098723P;
PR 02-SEP-1998; 98US-0098803P;
PR 02-SEP-1998; 98US-0098821P;
PR 02-SEP-1998; 98US-0098843P;
PR 09-SEP-1998; 98US-0099602P;
PR 10-SEP-1998; 98US-0099741P;
PR 10-SEP-1998; 98US-0099754P;
PR 10-SEP-1998; 98US-0099763P;
PR 10-SEP-1998; 98US-0099812P;
PR 15-SEP-1998; 98US-0100388P;
PR 16-SEP-1998; 98US-0100662P;
PR 16-SEP-1998; 98US-0100664P;
PR 16-SEP-1998; 98US-0101751P;
PR 16-SEP-1998; 98WO-US019330;
PR 17-SEP-1998; 98US-0100683P;
PR 17-SEP-1998; 98US-0100684P;
PR 17-SEP-1998; 98US-0100919P;
PR 17-SEP-1998; 98US-0100930P;
PR 18-SEP-1998; 98US-0100849P;
PR 18-SEP-1998; 98US-0101044P;
PR 18-SEP-1998; 98US-0101068P;
PR 23-SEP-1998; 98US-0101471P;
PR 23-SEP-1998; 98US-0101472P;
PR 23-SEP-1998; 98US-0101475P;
PR 23-SEP-1998; 98US-0101477P;
PR 24-SEP-1998; 98US-0101738P;
PR 24-SEP-1998; 98US-0101739P;
PR 24-SEP-1998; 98US-0101743P;

PR 24-SEP-1998; 98US-0101922P;
PR 25-SEP-1998; 98US-0101786P;
PR 29-SEP-1998; 98US-0102207P;
PR 29-SEP-1998; 98US-0102240P;
PR 29-SEP-1998; 98US-0102330P;
PR 29-SEP-1998; 98US-0102331P;
PR 30-SEP-1998; 98US-0102487P;
PR 30-SEP-1998; 98US-0102570P;
PR 30-SEP-1998; 98US-0102571P;
PR 01-OCT-1998; 98US-0102684P;
PR 01-OCT-1998; 98US-0102687P;

Query Match 100.0%; Score 1505; DB 7; Length 1505;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGGATCGGACCCCAAGCAGGTCCGGCGCGCGCGGAGAGCGCGCGCGGTTCAGCTCC 60
Db 1 CGCGGATCGGACCCCAAGCAGGTCCGGCGCGCGCGGAGAGCGCGCGCGGTTCAGCTCC 60

QY 61 TCGACCCCGGTGTCGGGCTAGTCCAGCGAGCGGCGCGGTGGGCCCATGGCCAGG 120
Db 61 TCGACCCCGGTGTCGGGCTAGTCCAGCGAGCGGCGCGGTGGGCCCATGGCCAGG 120

QY 121 CCGGCGATGGAGCGGTGGCGGACCGGCTGGCTGGTGACGGGGGCGCTCGGGGGGATC 180
Db 121 CCGGCGATGGAGCGGTGGCGGACCGGCTGGCTGGTGACGGGGGCGCTCGGGGGGATC 180

QY 181 GCGCGCGCGGTGGCGCGCGGCGCTGGTCCAGCAGGAGCTGAAAGGTGGTGGGCGCCCGC 240
Db 181 GCGCGCGCGGTGGCGCGCGGCGCTGGTCCAGCAGGAGCTGAAAGGTGGTGGGCGCGC 240

QY 241 ACTGTGGGCAACATCGAGGAGCTGGCTGCTGAATGTAAGATGAGGCTACCCGGGACT 300
Db 241 ACTGTGGGCAACATCGAGGAGCTGGCTGCTGAATGTAAGATGAGGCTACCCGGGACT 300

QY 301 TTGATCCCTACAGATGTGACCTATCAAAATGAAGGAGCATCTCTCCATGTTCTCAGCT 360
Db 301 TTGATCCCTACAGATGTGACCTATCAAAATGAAGGAGCATCTCTCCATGTTCTCAGCT 360

QY 361 ATCCGTTTCTCAGCAGCGGTGTAGACATCTGCATCAACAAATGCTGGCTTGGCCGGGCT 420
Db 361 ATCCGTTTCTCAGCAGCGGTGTAGACATCTGCATCAACAAATGCTGGCTTGGCCGGGCT 420

QY 421 GACACCTGCTCTCAGGAGCAGCAGGTGGTGGAGGAGCATGTTCAATGTGAAGTGTG 480
Db 421 GACACCTGCTCTCAGGAGCAGCAGGTGGTGGAGGAGCATGTTCAATGTGAAGTGTG 480

QY 481 GCCCTCAGCATCTGCACACGGGAAGCCTTACCAGTCCATGAAGGAGCGGAAATGTGGACGAT 540
Db 481 GCCCTCAGCATCTGCACACGGGAAGCCTTACCAGTCCATGAAGGAGCGGAAATGTGGACGAT 540

QY 541 GGGCACATCAATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGTCTGTGACC 600
Db 541 GGGCACATCAATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGTCTGTGACC 600

QY 601 CACTTCTATAGTGCACCAAGTATGCGGTCACTGGCTGACAGAGGAGCTCAGGCAAGAG 660
Db 601 CACTTCTATAGTGCACCAAGTATGCGGTCACTGGCTGACAGAGGAGCTCAGGCAAGAG 660

QY 661 CTTGGGAGGCCAGACCCACATCCGAGCCACGTGCATCTCTCCAGGTGGTGGAGACA 720
Db 661 CTTGGGAGGCCAGACCCACATCCGAGCCACGTGCATCTCTCCAGGTGGTGGAGACA 720

QY 721 CAATTGCGCTTCAAACTCCACGACAAAGGACCCCTGAGAAGGAGCTGCCACCTATGAGCAA 780
Db 721 CAATTGCGCTTCAAACTCCACGACAAAGGACCCCTGAGAAGGAGCTGCCACCTATGAGCAA 780

QY 781 ATGAAGTGTCTCAAAACCGGAGGTGTCGCGAGGCTGTTATCTACGTCTCAGCACCCCC 840
Db 781 ATGAAGTGTCTCAAAACCGGAGGTGTCGCGAGGCTGTTATCTACGTCTCAGCACCCCC 840

QY 841 GCACACATCCAGATTGGAGACATCCAGATGAGGCCACGGAGCAGGTGACCTAGTACTG 900
```

Db	841	GCACATCCAGATGGAGACATCCAGATGAGGCCACGAGCAGGTGACCTAGTACTG 900	PR 17-OCT-1997;	97US-0062250P.
			PR 21-OCT-1997;	97US-0063486P.
Qy	901	TGGAGCTCCTCTTCCCTCCCAACCTTCATAGCTTGCCTCTCTGGAATTTAGG 960	PR 24-OCT-1997;	97US-0063120P.
			PR 28-OCT-1997;	97US-0063121P.
Db	901	TGGAGCTCCTCTTCCCTCCCAACCTTCATAGCTTGCCTCTCTGGAATTTAGG 960	PR 28-OCT-1997;	97US-0063540P.
			PR 28-OCT-1997;	97US-0063541P.
Qy	961	TGTTGATTTCTGGATCAGGGATACCACTTCTCTGCACACCCCGACAGGGGTAGAAA 1020	PR 28-OCT-1997;	97US-0063544P.
			PR 28-OCT-1997;	97US-0063564P.
Db	961	TGTTGATTTCTGGATCAGGGATACCACTTCTCTGCACACCCCGACAGGGGTAGAAA 1020	PR 29-OCT-1997;	97US-0063734P.
			PR 31-OCT-1997;	97US-0063870P.
Qy	1021	ATTGTTTGAGATTTTATATCATCTTGTCAAATGTCTCAGTTGTAATGTGAAAAATG 1080	PR 31-OCT-1997;	97US-0064103P.
			PR 21-NOV-1997;	97US-0066120P.
Db	1021	ATTGTTTGAGATTTTATATCATCTTGTCAAATGTCTCAGTTGTAATGTGAAAAATG 1080	PR 24-NOV-1997;	97US-0066466P.
			PR 24-NOV-1997;	97US-0066772P.
Qy	1081	GGCTGGGGAAGAGGGTGGTCCCTAAATGTTTTTAACTGTTAACTGTTGTCGCCCC 1140	PR 11-DEC-1997;	97US-0069335P.
			PR 12-DEC-1997;	97US-0069425P.
Db	1081	GGCTGGGGAAGAGGGTGGTCCCTAAATGTTTTTAACTGTTAACTGTTGTCGCCCC 1140	PR 17-DEC-1997;	97US-0069870P.
			PR 18-DEC-1997;	97US-0068017P.
Qy	1141	TGGGCACTTGGCCTTCTGCTCCTCAGTGTCTTCCCTTTGACATGGGAAGGAGTTGTG 1200	PR 10-MAR-1998;	98US-0077450P.
			PR 11-MAR-1998;	98US-0077632P.
Db	1141	TGGGCACTTGGCCTTCTGCTCCTCAGTGTCTTCCCTTTGACATGGGAAGGAGTTGTG 1200	PR 11-MAR-1998;	98US-0077649P.
			PR 20-MAR-1998;	98US-0078886P.
Qy	1201	GCACAAATCCCATCTTCTTGACACTCAAAGTCTGTGGCTCAGGGCTGGGGTGGCAGAGG 1260	PR 20-MAR-1998;	98US-0078939P.
			PR 27-MAR-1998;	98US-0079664P.
Db	1201	GCACAAATCCCATCTTCTTGACACTCAAAGTCTGTGGCTCAGGGCTGGGGTGGCAGAGG 1260	PR 27-MAR-1998;	98US-0079786P.
			PR 31-MAR-1998;	98US-0080107P.
Qy	1261	GAGGCTTCACCTTATATCTGTGTGTTATCCAGGGCTCCAGACTTCTCTCTGCTGCTG 1320	PR 31-MAR-1998;	98US-0080194P.
			PR 01-APR-1998;	98US-0080327P.
Db	1261	GAGGCTTCACCTTATATCTGTGTGTTATCCAGGGCTCCAGACTTCTCTCTGCTGCTG 1320	PR 01-APR-1998;	98US-0080333P.
			PR 08-APR-1998;	98US-0081049P.
Qy	1321	CCCACTGCACCTCTCCCTTATCTATCTCTTCTCGGCTCCCGCCAGCTTCTGGCT 1380	PR 08-APR-1998;	98US-0081070P.
			PR 09-APR-1998;	98US-0081195P.
Db	1321	CCCACTGCACCTCTCCCTTATCTATCTCTTCTCGGCTCCCGCCAGCTTCTGGCT 1380	PR 15-APR-1998;	98US-0081838P.
			PR 21-APR-1998;	98US-0082568P.
Qy	1381	TCCTGTCCCTCTCGGGGTCACTCCCTCCACTCTGACTCTGACTATGGCAGCAGAACCA 1440	PR 21-APR-1998;	98US-0082569P.
			PR 22-APR-1998;	98US-0082704P.
Db	1381	TCCTGTCCCTCTCGGGGTCACTCCCTCCACTCTGACTCTGACTATGGCAGCAGAACCA 1440	PR 22-APR-1998;	98US-0082797P.
			PR 28-APR-1998;	98US-0083322P.
Qy	1441	GGGCTGGCCAGTGGATTTTCATGGTATCAATTAATAAGAAAAATCGAACCAAAAAA 1500	PR 29-APR-1998;	98US-0083496P.
			PR 29-APR-1998;	98US-0083499P.
Db	1441	GGGCTGGCCAGTGGATTTTCATGGTATCAATTAATAAGAAAAATCGAACCAAAAAA 1500	PR 29-APR-1998;	98US-0083559P.
			PR 05-MAY-1998;	98US-0084366P.
Qy	1501	AAAAA 1505	PR 06-MAY-1998;	98US-0084414P.
			PR 07-MAY-1998;	98US-0084639P.
Db	1501	AAAAA 1505	PR 07-MAY-1998;	98US-0084640P.
			PR 07-MAY-1998;	98US-0084643P.
Qy			PR 15-MAY-1998;	98US-0085579P.
			PR 15-MAY-1998;	98US-0085580P.
Db			PR 15-MAY-1998;	98US-0085582P.
			PR 18-MAY-1998;	98US-0086023P.
Qy			PR 22-MAY-1998;	98US-0086392P.
			PR 22-MAY-1998;	98US-0086486P.
Db			PR 28-MAY-1998;	98US-0087098P.
			PR 28-MAY-1998;	98US-0087208P.
Qy			PR 02-JUN-1998;	98US-0087609P.
			PR 02-JUN-1998;	98US-0087759P.
Db			PR 03-JUN-1998;	98US-0087827P.
			PR 04-JUN-1998;	98US-0088025P.
Qy			PR 04-JUN-1998;	98US-0088028P.
			PR 04-JUN-1998;	98US-0088029P.
Db			PR 04-JUN-1998;	98US-0088033P.
			PR 04-JUN-1998;	98US-0088326P.
Qy			PR 05-JUN-1998;	98US-0088167P.
			PR 05-JUN-1998;	98US-0088202P.
Db			PR 05-JUN-1998;	98US-0088212P.
			PR 09-JUN-1998;	98US-0088217P.
Qy			PR 10-JUN-1998;	98US-0088655P.
			PR 10-JUN-1998;	98US-0088722P.
Db			PR 10-JUN-1998;	98US-0088738P.
			PR 10-JUN-1998;	98US-0088740P.

RESULT 13

ACA98329

ID ACA98329 standard; cDNA; 1505 BP.

XX ACA98329;

AC ACA98329;

XX 25-JUL-2003 (first entry)

DT 25-JUL-2003 (first entry)

XX Novel human secreted and transmembrane protein PRO1774 cDNA.

DE Human; secreted and transmembrane protein; PRO; transgenic animal;

XX Human; secreted and transmembrane protein; PRO; transgenic animal;

KW knockout; chromosome identification; tissue typing; tumour;

KW chondrocyte proliferation; chondrocyte differentiation;

KW tumor necrosis factor-alpha release stimulator; gene; ss.

XX Homo sapiens.

OS Homo sapiens.

XX US2003036144-A1.

PN 20-FEB-2003.

XX 01-JUL-2002; 2002US-00187601.

PD 18-SEP-1997; 97US-0059263P.

XX 18-SEP-1997; 97US-0059266P.

PR 18-SEP-1997; 97US-0059266P.

PR 18-SEP-1997; 97US-0059266P.

PR	10-JUN-1998;	98US-0088811P.	PR	16-SEP-1998;	98US-0100664P.
PR	10-JUN-1998;	98US-0088824P.	PR	16-SEP-1998;	98US-0101751P.
PR	10-JUN-1998;	98US-0088825P.	PR	16-SEP-1998;	98WO-US019330.
PR	10-JUN-1998;	98US-0088826P.	PR	17-SEP-1998;	98US-0100683P.
PR	11-JUN-1998;	98US-0088861P.	PR	17-SEP-1998;	98US-0100694P.
PR	11-JUN-1998;	98US-0088863P.	PR	17-SEP-1998;	98US-0100919P.
PR	11-JUN-1998;	98US-0088876P.	PR	17-SEP-1998;	98US-0100930P.
PR	12-JUN-1998;	98US-0089000P.	PR	18-SEP-1998;	98US-0100849P.
PR	12-JUN-1998;	98US-0089105P.	PR	18-SEP-1998;	98US-0101014P.
PR	16-JUN-1998;	98US-0089512P.	PR	18-SEP-1998;	98US-0101068P.
PR	16-JUN-1998;	98US-0089514P.	PR	23-SEP-1998;	98US-0101471P.
PR	17-JUN-1998;	98US-0089538P.	PR	23-SEP-1998;	98US-0101472P.
PR	17-JUN-1998;	98US-0089539P.	PR	23-SEP-1998;	98US-0101475P.
PR	17-JUN-1998;	98US-0089653P.	PR	23-SEP-1998;	98US-0101477P.
PR	18-JUN-1998;	98US-0089908P.	PR	24-SEP-1998;	98US-0101738P.
PR	19-JUN-1998;	98US-0089952P.	PR	24-SEP-1998;	98US-0101739P.
PR	22-JUN-1998;	98US-0090246P.	PR	24-SEP-1998;	98US-0101743P.
PR	22-JUN-1998;	98US-0090252P.	PR	24-SEP-1998;	98US-0101922P.
PR	22-JUN-1998;	98US-0090254P.	PR	25-SEP-1998;	98US-0101786P.
PR	24-JUN-1998;	98US-0090429P.	PR	29-SEP-1998;	98US-0102207P.
PR	24-JUN-1998;	98US-0090435P.	PR	29-SEP-1998;	98US-0102240P.
PR	24-JUN-1998;	98US-0090444P.	PR	29-SEP-1998;	98US-0102330P.
PR	24-JUN-1998;	98US-0090461P.	PR	29-SEP-1998;	98US-0102331P.
PR	24-JUN-1998;	98US-0090535P.	PR	30-SEP-1998;	98US-0102487P.
PR	24-JUN-1998;	98US-0090540P.	PR	30-SEP-1998;	98US-0102570P.
PR	25-JUN-1998;	98US-0090676P.	PR	30-SEP-1998;	98US-0102571P.
PR	25-JUN-1998;	98US-0090678P.	PR	01-OCT-1998;	98US-0102684P.
PR	25-JUN-1998;	98US-0090688P.	PR	01-OCT-1998;	98US-0102687P.
PR	25-JUN-1998;	98US-0090690P.	PR	02-OCT-1998;	98US-0102965P.
PR	25-JUN-1998;	98US-0090695P.	PR	06-OCT-1998;	98US-0103258P.
PR	25-JUN-1998;	98US-0090696P.	PR	06-OCT-1998;	98US-0103449P.
PR	26-JUN-1998;	98US-00105413.	Query Match 100.0%; Score 1505; DB 7; Length 1505;		
PR	26-JUN-1998;	98US-0090862P.	Best Local Similarity 100.0%; Pred. No. 0;		
PR	26-JUN-1998;	98US-0090863P.	Matches 1505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
PR	26-JUN-1998;	98US-0091010P.	Qy	1	CGGGATCGGACCAAGCAGGTGGCGGCGCGGCGGAGAGCGCGGGCGGTGAGCTCC 60
PR	01-JUL-1998;	98US-0091359P.	Db	1	CGGGATCGGACCAAGCAGGTGGCGGCGCGGCGGAGAGCGCGGGCGGTGAGCTCC 60
PR	01-JUL-1998;	98US-0091544P.	Qy	61	TCGACCCCGGTGTCGGGCTAGTCCAGCGAGCGGCGGCGGCGGTGGGCCCATGGCCAGG 120
PR	02-JUL-1998;	98US-0091478P.	Db	61	TCGACCCCGGTGTCGGGCTAGTCCAGCGAGCGGCGGCGGCGGTGGGCCCATGGCCAGG 120
PR	02-JUL-1998;	98US-0091486P.	Qy	121	CCGGCATGAGAGCGGTGGCGGAGCCGGCTGGCGTGGTGTGACCGGGGGGCTCGGGGGGATC 180
PR	02-JUL-1998;	98US-0091628P.	Db	121	CCGGCATGAGAGCGGTGGCGGAGCCGGCTGGCGTGGTGTGACCGGGGGGCTCGGGGGGATC 180
PR	02-JUL-1998;	98US-0091632P.	Qy	181	GGCGCGCGGTGGCGCGGCGGCGCTGGTCCAGCAGGAGCTGAAGGTGGTGGGCTGCGCCGC 240
PR	04-AUG-1998;	98US-0094006P.	Db	181	GGCGCGCGGTGGCGCGGCGGCGCTGGTCCAGCAGGAGCTGAAGGTGGTGGGCTGCGCCGC 240
PR	04-AUG-1998;	98US-0095282P.	Qy	241	ACTGTGGGCAACATCGAGGAGTGGCTGTGAATGTAAAGTGCAGGCTACCCCGGGACT 300
PR	10-AUG-1998;	98US-0095938P.	Db	241	ACTGTGGGCAACATCGAGGAGTGGCTGTGAATGTAAAGTGCAGGCTACCCCGGGACT 300
PR	17-AUG-1998;	98US-0096757P.	Qy	301	TTGATCCCTACAGATGTGACCTTCAATGAAGAGGACATCCCTCTCATGTTCTCAGCT 360
PR	17-AUG-1998;	98US-0096766P.	Db	301	TTGATCCCTACAGATGTGACCTTCAATGAAGAGGACATCCCTCTCATGTTCTCAGCT 360
PR	17-AUG-1998;	98US-0096891P.	Qy	361	ATCGTTTCTCAGCACACGGGTGTAGACATCTGCATCAACAATGTGGCTTGGCCCGGCT 420
PR	17-AUG-1998;	98US-0096897P.	Db	361	ATCGTTTCTCAGCACACGGGTGTAGACATCTGCATCAACAATGTGGCTTGGCCCGGCT 420
PR	18-AUG-1998;	98US-0096949P.	Qy	421	GACACCTGTCTTCAGGCAGCACCGAGTGGTTGGAGGACATGTTCAATGTGAAGTCTGTG 480
PR	18-AUG-1998;	98US-0097022P.	Db	421	GACACCTGTCTTCAGGCAGCACCGAGTGGTTGGAGGACATGTTCAATGTGAAGTCTGTG 480
PR	26-AUG-1998;	98US-0097952P.	Qy	481	GCCCTCAGCATCTGCACACGGGAAGCCTACAGTCCATGAAGAGCGGAATGTGACGAT 540
PR	26-AUG-1998;	98US-0097954P.	Db	481	GCCCTCAGCATCTGCACACGGGAAGCCTACAGTCCATGAAGAGCGGAATGTGACGAT 540
PR	26-AUG-1998;	98US-0097955P.	Qy	541	GGGCACATCATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGTCTGTGACC 600
PR	26-AUG-1998;	98US-0097971P.	Db		
PR	26-AUG-1998;	98US-0098014P.	Qy		
PR	01-SEP-1998;	98US-0098716P.	Db		
PR	01-SEP-1998;	98US-0098723P.	Qy		
PR	02-SEP-1998;	98US-0098803P.	Db		
PR	02-SEP-1998;	98US-0098821P.	Qy		
PR	02-SEP-1998;	98US-0098843P.	Db		
PR	09-SEP-1998;	98US-0099602P.	Qy		
PR	10-SEP-1998;	98US-0099741P.	Db		
PR	10-SEP-1998;	98US-0099754P.	Qy		
PR	10-SEP-1998;	98US-0099763P.	Db		
PR	10-SEP-1998;	98US-0099812P.	Qy		
PR	15-SEP-1998;	98US-0100388P.	Db		
PR	16-SEP-1998;	98US-0100662P.	Qy		

Db 541 GGGCACATCAATAACATCAATAGATGTCGGCACCCGAGTGTACCCCTGCTGTGACC 600
QY 601 CACTTCTATAGTGCCACCAAGTAGTCCTGCTACTGCTGTGACAGAGGGAGTCAAGCAAGAG 660
Db 601 CACTTCTATAGTGCCACCAAGTAGTCCTGCTACTGCTGTGACAGAGGGAGTCAAGCAAGAG 660
QY 661 CTTGGGAGGCCAGACCCATCCGAGCAGGTGATCTCTCCAGGTGTGGTGAGACA 720
Db 661 CTTGGGAGGCCAGACCCATCCGAGCAGGTGATCTCTCCAGGTGTGGTGAGACA 720
QY 721 CAATTGCGCTTCAAACTCCAGCAAGACCCCTGAGAAGGCGAGCTGCCACCTATGAGCAA 780
Db 721 CAATTGCGCTTCAAACTCCAGCAAGACCCCTGAGAAGGCGAGCTGCCACCTATGAGCAA 780
QY 781 ATGAAGTGTCTCAAAACCGAGATGTGGCCGAGGCTGTTATCTAGCTCCTCAGCACCCCC 840
Db 781 ATGAAGTGTCTCAAAACCGAGATGTGGCCGAGGCTGTTATCTAGCTCCTCAGCACCCCC 840
QY 841 GCACATCCAGATGGAGACATCCAGATGAGGCCACGAGCAGGTGACCTAGTACTG 900
Db 841 GCACATCCAGATGGAGACATCCAGATGAGGCCACGAGCAGGTGACCTAGTACTG 900
QY 901 TGGGAGCTCCCTCCCTCCCCACCCCTTCATGGCTTGCCCTCCTGCCCTCTGGATTTAGG 960
Db 901 TGGGAGCTCCCTCCCTCCCCACCCCTTCATGGCTTGCCCTCCTGCCCTCTGGATTTAGG 960
QY 961 TGTGATTTCTGGATTCAGGGATACCACTTCCTGTGCAACCCCGACAGGGGTAGAAA 1020
Db 961 TGTGATTTCTGGATTCAGGGATACCACTTCCTGTGCAACCCCGACAGGGGTAGAAA 1020
QY 1021 ATTGTTTGGATTTTATATCATCTGTGCAAAATGCTTCAGTTGTAATGTGAAAAATG 1080
Db 1021 ATTGTTTGGATTTTATATCATCTGTGCAAAATGCTTCAGTTGTAATGTGAAAAATG 1080
QY 1081 GGCTGGGAAAGAGGTTGGTGTCCCTAATGTTTACTGTGTAATCTGTTGTCGCC 1140
Db 1081 GGCTGGGAAAGAGGTTGGTGTCCCTAATGTTTACTGTGTAATCTGTTGTCGCC 1140
QY 1141 TGGGACATGCGCCTTGTCTGCTCTCAGTGTCTTCCCTTTGACATGGAAAGGAGTTGTG 1200
Db 1141 TGGGACATGCGCCTTGTCTGCTCTCAGTGTCTTCCCTTTGACATGGAAAGGAGTTGTG 1200
QY 1201 GCCAAATCCCACTTCTTGCACCTCAACGCTGTGGCTCAGGCTGGGGTGGCAGAGG 1260
Db 1201 GCCAAATCCCACTTCTTGCACCTCAACGCTGTGGCTCAGGCTGGGGTGGCAGAGG 1260
QY 1261 GAGGCTTACCTTATATCTGTGTTGTTATCCAGGGCTCCAGACTTCTCTCTGCTGCTGC 1320
Db 1261 GAGGCTTACCTTATATCTGTGTTGTTATCCAGGGCTCCAGACTTCTCTCTGCTGCTGC 1320
QY 1321 CCAGTGCACCTCTCCCTCTTATCTATCTCTTCTGGCTCCCGAGCCAGTCTTGGCT 1380
Db 1321 CCAGTGCACCTCTCCCTCTTATCTATCTCTTCTGGCTCCCGAGCCAGTCTTGGCT 1380
QY 1381 TCTTGTCCCTCTGGGTGATCCCTCCACCTGACTCTGACTGATGGCAGCAGAACCA 1440
Db 1381 TCTTGTCCCTCTGGGTGATCCCTCCACCTGACTCTGACTGATGGCAGCAGAACCA 1440
QY 1441 GGGCCTGGCCAGTGGATTTTATGGTGATCATTAATAAGAAAAATCGCAACCAAAAAA 1500
Db 1441 GGGCCTGGCCAGTGGATTTTATGGTGATCATTAATAAGAAAAATCGCAACCAAAAAA 1500
QY 1501 AAAAA 1505
Db 1501 AAAAA 1505

RESULT 14
ACA93971
ID ACA93971 standard; cDNA; 1505 BP.
XX
AC ACA93971;

XX 18-JUL-2003 (first entry)
XX Human secreted/transmembrane protein (PRO) cDNA #189.
XX Human; ss; gene; secreted protein; transmembrane protein; PRO; tumour;
XX proliferation; differentiation; chondrocyte cell; TNF-alpha;
XX tumour necrosis factor-alpha; gene therapy.
XX Homo sapiens.
XX OS
XX US2003036149-A1.
XX 20-FEB-2003.
XX 02-JUL-2002; 2002US-00187746.
XX 18-SEP-1997; 97US-0059263P.
XX 18-SEP-1997; 97US-0059266P.
XX 17-OCT-1997; 97US-0062250P.
XX 21-OCT-1997; 97US-0063486P.
XX 24-OCT-1997; 97US-0063120P.
XX 24-OCT-1997; 97US-0063121P.
XX 28-OCT-1997; 97US-0063540P.
XX 28-OCT-1997; 97US-0063541P.
XX 28-OCT-1997; 97US-0063544P.
XX 28-OCT-1997; 97US-0063564P.
XX 29-OCT-1997; 97US-0063734P.
XX 31-OCT-1997; 97US-0063870P.
XX 31-OCT-1997; 97US-0064103P.
XX 13-NOV-1997; 97US-0065311P.
XX 21-NOV-1997; 97US-0066120P.
XX 24-NOV-1997; 97US-0066466P.
XX 24-NOV-1997; 97US-0066772P.
XX 11-DEC-1997; 97US-0069335P.
XX 12-DEC-1997; 97US-0069425P.
XX 17-DEC-1997; 97US-0069870P.
XX 18-DEC-1997; 97US-0068017P.
XX 10-MAR-1998; 98US-0077450P.
XX 11-MAR-1998; 98US-0077632P.
XX 11-MAR-1998; 98US-0077649P.
XX 20-MAR-1998; 98US-0078866P.
XX 20-MAR-1998; 98US-0078939P.
XX 27-MAR-1998; 98US-0079664P.
XX 27-MAR-1998; 98US-0079786P.
XX 31-MAR-1998; 98US-0080107P.
XX 31-MAR-1998; 98US-0080194P.
XX 01-APR-1998; 98US-0080327P.
XX 01-APR-1998; 98US-0080333P.
XX 08-APR-1998; 98US-0081049P.
XX 08-APR-1998; 98US-0081070P.
XX 09-APR-1998; 98US-0081195P.
XX 15-APR-1998; 98US-0081838P.
XX 21-APR-1998; 98US-0082568P.
XX 21-APR-1998; 98US-0082569P.
XX 22-APR-1998; 98US-0082704P.
XX 22-APR-1998; 98US-0082797P.
XX 28-APR-1998; 98US-0083322P.
XX 29-APR-1998; 98US-0083495P.
XX 29-APR-1998; 98US-0083496P.
XX 29-APR-1998; 98US-0083499P.
XX 29-APR-1998; 98US-0083559P.
XX 05-MAY-1998; 98US-0084366P.
XX 06-MAY-1998; 98US-0084414P.
XX 07-MAY-1998; 98US-0084639P.
XX 07-MAY-1998; 98US-0084640P.
XX 07-MAY-1998; 98US-0084643P.
XX 15-MAY-1998; 98US-0085579P.
XX 15-MAY-1998; 98US-0085580P.
XX 15-MAY-1998; 98US-0085582P.
XX 15-MAY-1998; 98US-0085700P.
XX 18-MAY-1998; 98US-0086023P.
XX 22-MAY-1998; 98US-0086392P.

PR	22-MAY-1998;	98US-0086486P.	PR	18-AUG-1998;	98US-0097022P.
PR	28-MAY-1998;	98US-0087098P.	PR	26-AUG-1998;	98US-0097952P.
PR	02-MAY-1998;	98US-0087208P.	PR	26-AUG-1998;	98US-0097954P.
PR	02-JUN-1998;	98US-0087539P.	PR	26-AUG-1998;	98US-0097955P.
PR	03-JUN-1998;	98US-0087827P.	PR	26-AUG-1998;	98US-0097971P.
PR	04-JUN-1998;	98US-0088025P.	PR	26-AUG-1998;	98US-0097974P.
PR	04-JUN-1998;	98US-0088028P.	PR	01-SEP-1998;	98US-0098014P.
PR	04-JUN-1998;	98US-0088029P.	PR	01-SEP-1998;	98US-0098716P.
PR	04-JUN-1998;	98US-0088033P.	PR	02-SEP-1998;	98US-0098723P.
PR	04-JUN-1998;	98US-0088326P.	PR	02-SEP-1998;	98US-0098803P.
PR	05-JUN-1998;	98US-0088167P.	PR	02-SEP-1998;	98US-0098821P.
PR	05-JUN-1998;	98US-0088202P.	PR	09-SEP-1998;	98US-0098843P.
PR	05-JUN-1998;	98US-0088212P.	PR	10-SEP-1998;	98US-0099602P.
PR	05-JUN-1998;	98US-0088217P.	PR	10-SEP-1998;	98US-0099741P.
PR	09-JUN-1998;	98US-0088655P.	PR	10-SEP-1998;	98US-0099754P.
PR	10-JUN-1998;	98US-0088722P.	PR	10-SEP-1998;	98US-0099763P.
PR	10-JUN-1998;	98US-0088738P.	PR	15-SEP-1998;	98US-0099812P.
PR	10-JUN-1998;	98US-0088740P.	PR	15-SEP-1998;	98US-0100388P.
PR	10-JUN-1998;	98US-0088811P.	PR	16-SEP-1998;	98US-0100662P.
PR	10-JUN-1998;	98US-0088824P.	PR	16-SEP-1998;	98US-0100664P.
PR	10-JUN-1998;	98US-0088825P.	PR	16-SEP-1998;	98US-0101751P.
PR	10-JUN-1998;	98US-0088826P.	PR	16-SEP-1998;	98US-0101751P.
PR	11-JUN-1998;	98US-0088861P.	PR	17-SEP-1998;	98US-0100683P.
PR	11-JUN-1998;	98US-0088863P.	PR	17-SEP-1998;	98US-0100684P.
PR	11-JUN-1998;	98US-0088876P.	PR	17-SEP-1998;	98US-0100919P.
PR	12-JUN-1998;	98US-0089090P.	PR	17-SEP-1998;	98US-0100930P.
PR	12-JUN-1998;	98US-0089103P.	PR	18-SEP-1998;	98US-0100849P.
PR	16-JUN-1998;	98US-0089512P.	PR	18-SEP-1998;	98US-0101014P.
PR	16-JUN-1998;	98US-0089514P.	PR	18-SEP-1998;	98US-0101068P.
PR	17-JUN-1998;	98US-0089538P.	PR	23-SEP-1998;	98US-0101471P.
PR	17-JUN-1998;	98US-0089598P.	PR	23-SEP-1998;	98US-0101472P.
PR	17-JUN-1998;	98US-0089653P.	PR	23-SEP-1998;	98US-0101475P.
PR	18-JUN-1998;	98US-0089908P.	PR	23-SEP-1998;	98US-0101477P.
PR	19-JUN-1998;	98US-0089952P.	PR	24-SEP-1998;	98US-0101738P.
PR	22-JUN-1998;	98US-0090246P.	PR	24-SEP-1998;	98US-0101739P.
PR	22-JUN-1998;	98US-0090252P.	PR	24-SEP-1998;	98US-0101743P.
PR	22-JUN-1998;	98US-0090254P.	PR	25-SEP-1998;	98US-0101922P.
PR	24-JUN-1998;	98US-0090423P.	PR	25-SEP-1998;	98US-0101786P.
PR	24-JUN-1998;	98US-0090435P.	PR	29-SEP-1998;	98US-0102207P.
PR	24-JUN-1998;	98US-0090444P.	PR	29-SEP-1998;	98US-0102240P.
PR	24-JUN-1998;	98US-0090461P.	PR	29-SEP-1998;	98US-0102330P.
PR	24-JUN-1998;	98US-0090535P.	PR	29-SEP-1998;	98US-0102331P.
PR	24-JUN-1998;	98US-0090540P.	PR	30-SEP-1998;	98US-0102487P.
PR	25-JUN-1998;	98US-0090676P.	PR	30-SEP-1998;	98US-0102570P.
PR	25-JUN-1998;	98US-0090678P.	PR	30-SEP-1998;	98US-0102571P.
PR	25-JUN-1998;	98US-0090688P.	PR	01-OCT-1998;	98US-0102684P.
PR	25-JUN-1998;	98US-0090690P.	PR	01-OCT-1998;	98US-0102687P.
PR	25-JUN-1998;	98US-0090694P.	PR	02-OCT-1998;	98US-0102965P.
PR	25-JUN-1998;	98US-0090695P.	PR	06-OCT-1998;	98US-0103258P.
PR	25-JUN-1998;	98US-0090696P.	PR	06-OCT-1998;	98US-0103449P.
PR	26-JUN-1998;	98US-00105413.	PR	07-OCT-1998;	98US-00168978.
PR	26-JUN-1998;	98US-0090862P.			
PR	26-JUN-1998;	98US-0090863P.			
PR	26-JUN-1998;	98US-0091010P.			
PR	01-JUL-1998;	98US-0091359P.			
PR	01-JUL-1998;	98US-0091544P.			
PR	02-JUL-1998;	98US-0091478P.			
PR	02-JUL-1998;	98US-0091486P.			
PR	02-JUL-1998;	98US-0091626P.			
PR	02-JUL-1998;	98US-0091628P.			
PR	02-JUL-1998;	98US-0091632P.			
PR	24-JUL-1998;	98US-0094006P.			
PR	04-AUG-1998;	98US-0095282P.			
PR	10-AUG-1998;	98US-0095998P.			
PR	10-AUG-1998;	98US-0096012P.			
PR	17-AUG-1998;	98US-0096757P.			
PR	17-AUG-1998;	98US-0096766P.			
PR	17-AUG-1998;	98US-0096867P.			
PR	17-AUG-1998;	98US-0096891P.			
PR	17-AUG-1998;	98US-0096897P.			
PR	18-AUG-1998;	98US-0096949P.			
PR	18-AUG-1998;	98US-0096959P.			

Query Match

Best Local Similarity

Matches 1505; Conservative

100.0%;

100.0%;

0; Mismatches

0; Indels

0; Gaps

0;

Qy	1	CGCGGATCGGACCCAGCAGGTTCGGCGCGCGCGGCGGAGAGCGCGCGCGGTTCAGCTCC	60
Db	1	CGCGGATCGGACCCAGCAGGTTCGGCGCGCGCGGCGGAGAGCGCGCGCGGTTCAGCTCC	60
Qy	61	TCGACCCCGGTGTTCGGGCTAGTCCAGCGAGCGGCGCGCGGTGGGCCCATGGCCAGG	120
Db	61	TCGACCCCGGTGTTCGGGCTAGTCCAGCGAGCGGCGCGCGGTGGGCCCATGGCCAGG	120
Qy	121	CCCGGATGAGCGGTGGCGGCGACCGGCTGGCGCTGTGTGACGGGGCCCTCGGGGGGCAATC	180
Db	121	CCCGGATGAGCGGTGGCGGCGACCGGCTGGCGCTGTGTGACGGGGCCCTCGGGGGGCAATC	180
Qy	181	GGCGGGCGGTGGCCCGGGCCCTGGTCCAGCAGGACTGAAAGGTGGTGGGCTGCCCCGCG	240
Db	181	GGCGGGCGGTGGCCCGGGCCCTGGTCCAGCAGGACTGAAAGGTGGTGGGCTGCCCCGCG	240
Qy	241	ACTGTGGGCAACATCGAGGAGTGGCTCTGTAATGTAAGAGTGCAGGCTACCCCGGACT	300

PR 15-APR-1998; 98US-0081838P.
 PR 21-APR-1998; 98US-0082568P.
 PR 21-APR-1998; 98US-0082569P.
 PR 22-APR-1998; 98US-0082704P.
 PR 22-APR-1998; 98US-0082797P.
 PR 28-APR-1998; 98US-0083322P.
 PR 29-APR-1998; 98US-0083495P.
 PR 29-APR-1998; 98US-0083496P.
 PR 29-APR-1998; 98US-0083499P.
 PR 29-APR-1998; 98US-0083559P.
 PR 05-MAY-1998; 98US-0084366P.
 PR 06-MAY-1998; 98US-0084414P.
 PR 07-MAY-1998; 98US-0084639P.
 PR 07-MAY-1998; 98US-0084640P.
 PR 07-MAY-1998; 98US-0084643P.
 PR 15-MAY-1998; 98US-0085579P.
 PR 15-MAY-1998; 98US-0085580P.
 PR 15-MAY-1998; 98US-0085582P.
 PR 15-MAY-1998; 98US-0085700P.
 PR 18-MAY-1998; 98US-0086023P.
 PR 22-MAY-1998; 98US-0086332P.
 PR 22-MAY-1998; 98US-0086486P.
 PR 28-MAY-1998; 98US-0087098P.
 PR 28-MAY-1998; 98US-0087208P.
 PR 02-JUN-1998; 98US-0087609P.
 PR 02-JUN-1998; 98US-0087759P.
 PR 03-JUN-1998; 98US-0087827P.
 PR 04-JUN-1998; 98US-0088025P.
 PR 04-JUN-1998; 98US-0088028P.
 PR 04-JUN-1998; 98US-0088029P.
 PR 04-JUN-1998; 98US-0088033P.
 PR 04-JUN-1998; 98US-0088326P.
 PR 05-JUN-1998; 98US-0088167P.
 PR 05-JUN-1998; 98US-0088202P.
 PR 05-JUN-1998; 98US-0088212P.
 PR 05-JUN-1998; 98US-0088217P.
 PR 09-JUN-1998; 98US-0088655P.
 PR 10-JUN-1998; 98US-0088732P.
 PR 10-JUN-1998; 98US-0088738P.
 PR 10-JUN-1998; 98US-0088740P.
 PR 10-JUN-1998; 98US-0088811P.
 PR 10-JUN-1998; 98US-0088824P.
 PR 10-JUN-1998; 98US-0088825P.
 PR 10-JUN-1998; 98US-0088836P.
 PR 11-JUN-1998; 98US-0088861P.
 PR 11-JUN-1998; 98US-0088863P.
 PR 11-JUN-1998; 98US-0088876P.
 PR 12-JUN-1998; 98US-0089030P.
 PR 12-JUN-1998; 98US-0089105P.
 PR 16-JUN-1998; 98US-0089512P.
 PR 16-JUN-1998; 98US-0089514P.
 PR 17-JUN-1998; 98US-0089538P.
 PR 17-JUN-1998; 98US-0089598P.
 PR 17-JUN-1998; 98US-0089653P.
 PR 18-JUN-1998; 98US-0089908P.
 PR 19-JUN-1998; 98US-0089952P.
 PR 22-JUN-1998; 98US-0090246P.
 PR 22-JUN-1998; 98US-0090252P.
 PR 22-JUN-1998; 98US-0090254P.
 PR 24-JUN-1998; 98US-0090429P.
 PR 24-JUN-1998; 98US-0090435P.
 PR 24-JUN-1998; 98US-0090444P.
 PR 24-JUN-1998; 98US-0090461P.
 PR 24-JUN-1998; 98US-0090535P.
 PR 24-JUN-1998; 98US-0090540P.
 PR 25-JUN-1998; 98US-0090678P.
 PR 25-JUN-1998; 98US-0090688P.
 PR 25-JUN-1998; 98US-0090690P.
 PR 25-JUN-1998; 98US-0090694P.
 PR 25-JUN-1998; 98US-0090695P.
 PR 26-JUN-1998; 98US-0090696P.
 PR 26-JUN-1998; 98US-00105413.

PR 26-JUN-1998; 98US-0090862P.
 PR 26-JUN-1998; 98US-0090863P.
 PR 26-JUN-1998; 98US-0091010P.
 PR 01-JUL-1998; 98US-0091359P.
 PR 01-JUL-1998; 98US-0091544P.
 PR 02-JUL-1998; 98US-0091478P.
 PR 02-JUL-1998; 98US-0091486P.
 PR 02-JUL-1998; 98US-0091626P.
 PR 02-JUL-1998; 98US-0091628P.
 PR 02-JUL-1998; 98US-0091632P.
 PR 04-AUG-1998; 98US-0094006P.
 PR 04-AUG-1998; 98US-0095282P.
 PR 10-AUG-1998; 98US-0095998P.
 PR 10-AUG-1998; 98US-0096012P.
 PR 17-AUG-1998; 98US-0096757P.
 PR 17-AUG-1998; 98US-0096766P.
 PR 17-AUG-1998; 98US-0096867P.
 PR 17-AUG-1998; 98US-0096891P.
 PR 17-AUG-1998; 98US-0096897P.
 PR 18-AUG-1998; 98US-0096949P.
 PR 18-AUG-1998; 98US-0096959P.
 PR 18-AUG-1998; 98US-0097022P.
 PR 26-AUG-1998; 98US-0097952P.
 PR 26-AUG-1998; 98US-0097954P.
 PR 26-AUG-1998; 98US-0097955P.
 PR 26-AUG-1998; 98US-0097971P.
 PR 26-AUG-1998; 98US-0097974P.
 PR 26-AUG-1998; 98US-0098014P.
 PR 01-SEP-1998; 98US-0098716P.
 PR 01-SEP-1998; 98US-0098723P.
 PR 02-SEP-1998; 98US-0098803P.
 PR 02-SEP-1998; 98US-0098821P.
 PR 02-SEP-1998; 98US-0098843P.
 PR 09-SEP-1998; 98US-0099602P.
 PR 10-SEP-1998; 98US-0099741P.
 PR 10-SEP-1998; 98US-0099754P.
 PR 10-SEP-1998; 98US-0099763P.
 PR 15-SEP-1998; 98US-0099812P.
 PR 16-SEP-1998; 98US-0100388P.
 PR 16-SEP-1998; 98US-0100662P.
 PR 16-SEP-1998; 98US-0100664P.
 PR 16-SEP-1998; 98US-0101751P.
 PR 16-SEP-1998; 98WO-US019330.
 PR 17-SEP-1998; 98US-0100683P.
 PR 17-SEP-1998; 98US-0100684P.
 PR 17-SEP-1998; 98US-0100919P.
 PR 17-SEP-1998; 98US-0100930P.
 PR 18-SEP-1998; 98US-0100849P.
 PR 18-SEP-1998; 98US-0101014P.
 PR 23-SEP-1998; 98US-0101068P.
 PR 23-SEP-1998; 98US-0101471P.
 PR 23-SEP-1998; 98US-0101472P.
 PR 23-SEP-1998; 98US-0101475P.
 PR 23-SEP-1998; 98US-0101477P.
 PR 24-SEP-1998; 98US-0101738P.
 PR 24-SEP-1998; 98US-0101739P.
 PR 24-SEP-1998; 98US-0101743P.
 PR 24-SEP-1998; 98US-0101922P.
 PR 25-SEP-1998; 98US-0101786P.
 PR 29-SEP-1998; 98US-0102207P.
 PR 29-SEP-1998; 98US-0102240P.
 PR 29-SEP-1998; 98US-0102330P.
 PR 29-SEP-1998; 98US-0102331P.
 PR 30-SEP-1998; 98US-0102487P.
 PR 30-SEP-1998; 98US-0102570P.
 PR 30-SEP-1998; 98US-0102571P.
 PR 01-OCT-1998; 98US-0102684P.
 PR 01-OCT-1998; 98US-0102687P.
 PR 02-OCT-1998; 98US-0102965P.
 PR 06-OCT-1998; 98US-0103258P.
 PR 06-OCT-1998; 98US-0103449P.

Query Match

100.0%; Score 1505; DB 7; Length 1505;

[illegible]

B/ack

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

```

OM nucleic - nucleic search, using sw model

Run on:      September 16, 2004, 12:21:25 ; Search time 5837 Seconds
              (without alignments)
              11175.478 Million cell updates/sec

Title:       US-10-063-735-127
Perfect score: 1505
Sequence:    1 gcgggatcggaaccgaagcag.....tcgcaaccacaaaaaaaaa 1505

Scoring table: OLIGO NUC
                Gapop 60.0 , Gapext 60.0

Searched:    3470272 seqs, 21671516995 residues

Word size :   10

Total number of hits satisfying chosen parameters: 3307169

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

```

[illegible]

ALIGNMENTS

[illegible]

Pred. No. is the number of results predicted by chance to have a

Query Match	100.0%;	Score 1505;	DB 6;	Length 1505;
Best Local Similarity	100.0%;	Prod. No. 0;		
Matches 1505;	Conservative	0;	Mismatches 0;	Indels 0;
Gaps	0;			
Qy	1	CGCGATCGGACCCAAAGCAGGTGCGCGCGCGGCGGAGGAGCGCGCGGGGTGAGCTCC	60	
Db	1	CGCGATCGGACCCAAAGCAGGTGCGCGCGCGGCGGAGGAGCGCGCGGGGTGAGCTCC	60	
Qy	61	TCGACCCCCCGTGTGCGGCTAGTCCAGCGAGGCGGACCGGCGGCGTGGGCCCATGCCAGG	120	
Db	61	TCGACCCCCCGTGTGCGGCTAGTCCAGCGAGGCGGACCGGCGGCGTGGGCCCATGCCAGG	120	
Qy	121	CCGGCATGGAGCGGTGGCGGACCGGCTGGCGTGTGAACGGGGGCTCGGGGGCATC	180	
Db	121	CCGGCATGGAGCGGTGGCGGACCGGCTGGCGTGTGAACGGGGGCTCGGGGGCATC	180	
Qy	181	GGCGGGCGGTGGCCGGCCCTGGTCCAGCAGGACCTGAAGTGGTGGCTGGCCCGCCG	240	
Db	181	GGCGGGCGGTGGCCGGCCCTGGTCCAGCAGGACCTGAAGTGGTGGCTGGCCCGCCG	240	
Qy	241	ACTGTGGCAACATCGAGGAGCTGGCTGCTGAAATGTAAGAGTGCAAGGTACCCCGGACT	300	
Db	241	ACTGTGGCAACATCGAGGAGCTGGCTGCTGAAATGTAAGAGTGCAAGGTACCCCGGACT	300	
Qy	301	TTGATCCCTACAGATGTGACCTATCAATGAAGAGGACATCTCTCCATGTCTCAGCT	360	
Db	301	TTGATCCCTACAGATGTGACCTATCAATGAAGAGGACATCTCTCCATGTCTCAGCT	360	
Qy	361	ATCCGTTCTCAGCACAGCGGTGAGACATCTGCATCAACATGCTGGCTGGCCGGCT	420	
Db	361	ATCCGTTCTCAGCACAGCGGTGAGACATCTGCATCAACATGCTGGCTGGCCGGCT	420	
Qy	421	GACACCTGTCTCTCAGGCAGCACAGTGGTTGGAGGACATGTTCAATGTCAAAGTGGTG	480	
Db	421	GACACCTGTCTCTCAGGCAGCACAGTGGTTGGAGGACATGTTCAATGTCAAAGTGGTG	480	
Qy	481	GCCCTCAGCATCTGCACAGGGAAGCTTACAGTCCATGAAGGAGCGGAATGTGACGAT	540	
Db	481	GCCCTCAGCATCTGCACAGGGAAGCTTACAGTCCATGAAGGAGCGGAATGTGACGAT	540	
Qy	541	GGGCACATCAATAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGTCTGTGACC	600	
Db	541	GGGCACATCAATAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGTCTGTGACC	600	
Qy	601	CACCTCTATAGTGCCACCAAGTATGCGGTCACTCGCTGACAGAGGAGCTGAGCAAGAG	660	
Db	601	CACCTCTATAGTGCCACCAAGTATGCGGTCACTCGGTGACAGAGGAGCTGAGCAAGAG	660	
Qy	661	CTTCGGAGGCGCAGACCAATCCGAGCCACGTGCATCTCTCAGGTGGTGGAGACA	720	
Db	661	CTTCGGAGGCGCAGACCAATCCGAGCCACGTGCATCTCTCAGGTGGTGGAGACA	720	
Qy	721	CAATTGCGCTTCAAACTCCAGCAAGGACCCCTGAGAGGAGCTGCCATATGAGGAA	780	
Db	721	CAATTGCGCTTCAAACTCCAGCAAGGACCCCTGAGAGGAGCTGCCATATGAGGAA	780	
Qy	781	ATGAAGTGTCTCAAAACCGAGGATGTGGCCGAGGCTGTTATCTACGTCTCAGCACCCCC	840	
Db	781	ATGAAGTGTCTCAAAACCGAGGATGTGGCCGAGGCTGTTATCTACGTCTCAGCACCCCC	840	
Qy	841	GCACATCCAGATTGGAGACATCCAGATGAGGCCCAACCGAGGAGTGAACCTAGTGACTG	900	
Db	841	GCACATCCAGATTGGAGACATCCAGATGAGGCCCAACCGAGGAGTGAACCTAGTGACTG	900	
Qy	901	TGGAGCTCTCCTTCCCTCCCAACCTTCATGGCTTGCCTCTGCCTCTGGAATTTAGG	960	
Db	901	TGGAGCTCTCCTTCCCTCCCAACCTTCATGGCTTGCCTCTGCCTCTGGAATTTAGG	960	
Qy	961	TGTTGATTTCTGGATCACGGGATACCACTTCTCTCCACACCCCGACCGGGGTAGAAA	1020	
Db	961	TGTTGATTTCTGGATCACGGGATACCACTTCTCTCCACACCCCGACCGGGGTAGAAA	1020	
Qy	1021	ATTGTTTGAGATTTTATATCATCTTGTCAAAATGTGCTTCAAGTTGTAATGTGAAAAATG	1080	

Db 1330 ATTGTTTGAGATTTTATATCATCTTGTCAAAATGCTTCAGTTGTAAATGTGAAAAATG 1389

QY 1081 GGCTGGGAAAGAGGTGGTCCCTAAATGTTTACTTGTAACTGTTAACTGTTCTGTCGCC 1140

Db 1390 GGTGGGAAAGAGGTGGTCCCTAAATGTTTACTTGTAACTGTTCTGTCGCC 1449

QY 1141 TGGGCATTTGGCCCTTTGCTGCTCTCAGTGTCTTCCCTTTGACATGGGAAAGAGTTGTG 1200

Db 1450 TGGGCATTTGGCCCTTTGCTGCTCTCAGTGTCTTCCCTTTGACATGGGAAAGAGTTGTG 1509

QY 1201 GCCAAATCCCATCTTCTTGACCTCAACGTCGTGTGGTCA-GGGCTGGGGTGGCAGAG 1259

Db 1510 GCCAAATCCCATCTTCTTGACCTCAACGTCGTGTGGTCAACGTCGTGGTGGCAGAG 1569

QY 1260 GGAGGCTTCACCTATATCTGTGTTGTTATCCAGGCTCCAGACTTCCTCTGCTGCTG 1319

Db 1570 GGAGGCTTCACCTATATCTGTGTTGTTATCCAGGCTCCAGACTTCCTCTGCTGCTG 1629

QY 1320 CCCCACTGCACCTCTCCCTTTATCTATCTCTCTCGGCTCCCGAGCCAGTCTTGGC 1379

Db 1630 CCCCACTGCACCTCTCCCTTTATCTATCTCTCTCGGCTCCCGAGCCAGTCTTGGC 1689

QY 1380 TTCCTGTCCTCTCGGGGTCAATCCCTCACTCTGACTCTGACTATGGCAGCAGAACACC 1439

Db 1690 TTCCTGTCCTCTCGGGGTCAATCCCTCACTCTGACTCTGACTATGGCAGCAGAACACC 1749

QY 1440 AGGGCTGCGCCAGTGGATTTCATGCTGATCAATTAAGAAAGAAATCCGACCAAAAAA 1499

Db 1750 AGGGCTGCGCCAGTGGATTTCATGCTGATCAATTAAGAAAGAAATCCGACCAAAAAA 1809

QY 1500 AAAAAA 1505

Db 1810 AAAAAA 1815

RESULT 6

AK026196

LOCUS Homo sapiens cDNA: FLJ22543 fis, clone HSI00212. linear PRI 13-SEP-2003

DEFINITION AK026196

ACCESSION AK026196

VERSION AK026196.1 GI:10438967

KEYWORDS oligo capping; fis (full insert sequence).

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1

AUTHORS Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.

TITLE NEDO human cDNA sequencing project

JOURNAL Unpublished

AUTHORS 2 (bases 1 to 1410)

Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T., Shibahara,T., Tanaka,T. and Nakamura,Y.

Direct Submission

Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail: flicdn@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)

COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'- and 3'-end one pass sequencing; Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).

FEATURES

Location/Qualifiers

1..1410

organism="Homo sapiens"

mol_type="mRNA"

/db_xref="taxon:9606"

/clone="HSI00212"

/tissue_type="human small intestine"

/clone_fib="HSI"

/note="cloning vector pME18SFL3"

243..788

/note="unnamed protein product"

/codon_start=1

/protein_id="BAB15390.1"

/db_xref="GI:10438968"

/translation="MSSAIRSQSHGVDCINNAGLARPTLLSGSTSGWKDMFNVL

ALSICTRAYSOMKERNVDDGHIININMSGHRVLPVSVTHEYSAKVAATYTBGLR

CELREAQTHIRATCISPGVETQPAFKLHDXDPKAAATYEQMKLCKPDAEVAEVIIV

LSTPAHIQIGDIQMRPTEQVI"

ORIGIN

Query Match 80.3%; Score 1209; DB 9; Length 1410;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 246 GGGCAACATCGAGGAGCTGGCTGCTGAATGTAAGAGTCAGGCTACCCCGGACTTTGAT 305

Db 140 GGGCAACATCGAGGAGCTGGCTGCTGAATGTAAGAGTCAGGCTACCCCGGACTTTGAT 199

QY 306 CCCCTACAGATGTGACCTTCAATGAAGAGACATCTCTCCATGTTCTAGCTATCGG 365

Db 200 CCCCTACAGATGTGACCTTCAATGAAGAGACATCTCTCCATGTTCTAGCTATCGG 259

QY 366 TTCTCAGCACAGCGGTGTAGACATCTGCATCAACATGCTGGCTTGGCCGCGCTGACAC 425

Db 260 TTCTCAGCACAGCGGTGTAGACATCTGCATCAACATGCTGGCTTGGCCGCGCTGACAC 319

QY 426 CCTGCTCTCAGGACGACCATGCTGTTGGAAGACATGTTCAATGTGAACGTTCTGCGCCCT 485

Db 320 CCTGCTCTCAGGACGACCATGCTGTTGGAAGACATGTTCAATGTGAACGTTCTGCGCCCT 379

QY 486 CAGCATCTGCACACGGGAAGCTTACCAGTCCATGAAGAGCGGAATGTGGACATGGGCA 545

Db 380 CAGCATCTGCACACGGGAAGCTTACCAGTCCATGAAGAGCGGAATGTGGACATGGGCA 439

QY 546 CATCATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGTCTGTGACCCACTT 605

Db 440 CATCATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGTCTGTGACCCACTT 499

QY 606 CTATAGTCCACCAAGTATGCGCTGCTGACAGGAGCTGAGGCAAGAGCTTCG 665

Db 500 CTATAGTCCACCAAGTATGCGCTGCTGACAGGAGCTGAGGCAAGAGCTTCG 559

QY 666 GGAGGCCAGACCCACATCCGAGCCAGTGCATCTCTCCAGTGTGTGGAGACACATTT 725

Db 560 GGAGGCCAGACCCACATCCGAGCCAGTGCATCTCTCCAGTGTGTGGAGACACATTT 619

QY 726 CGCCTTCAAACTCCACGACAGGACCCCTGAGAAGCGAGCTGCCACCTATAGCAAAATGA 785

Db 620 CGCCTTCAAACTCCACGACAGGACCCCTGAGAAGCGAGCTGCCACCTATAGCAAAATGA 679

QY 786 GTGTCTCAAACTCCAGAGTGTGGCCGAGGTGTTATCTACGCTCTCAGACACCCCGCACA 845

Db 680 GTGTCTCAAACTCCAGAGTGTGGCCGAGGTGTTATCTACGCTCTCAGACACCCCGCACA 739

QY 846 CATCCAGATTGGAGACATCCAGATGAGGCCACCGAGCAGGTGACTAGTACTGTGGGA 905

Db 740 CATCCAGATTGGAGACATCCAGATGAGGCCACCGAGCAGGTGACTAGTACTGTGGGA 799

QY 906 GCTCCTCTTCCCTCCCGACCTTTCATGGCTTGCTCTGCTCTGAGTTTATGAGTGTG 965

Db 800 GCTCCTCTTCCCTCCCGACCTTTCATGGCTTGCTCTGCTCTGAGTTTATGAGTGTG 859

QY 966 ATTTCTGGATCAGGGATACACTTCTGCTGCTTCCACACCCCGACCGAGGGCTAGAAAAATTTG 1025

Db 860 ATTTCTGGATCAGGGATACACTTCTGCTGCTTCCACACCCCGACCGAGGGCTAGAAAAATTTG 919

QY 1026 TTTTGGATTTTATATCATCTTGTCAAAATGCTTCAATGTTGAAATGTGAAAAATTTGCGGCTG 1085

Db 920 TTTGAGATTTTATATCAATCTGTCMAATGCTTCAAGTTGTAATCTGTAATCTGAAATATGGGCTG 979

Qy 1086 GGGAAAGAGGTGTGTGCTCCCTAAATGTTTAACTTTACTTTGTTAACTTGTCTTGTGCCCTGGGC 1145

Db 980 GGGAAAGAGGTGTGTGCTCCCTAAATGTTTAACTTTACTTTGTTAACTTGTCTTGTGCCCTGGGC 1039

Qy 1446 ACTTGGCCCTTGTCTGCTCTAGTGTCTTCCCTTGAATGGGAAGAGTGTGGCCAA 1205

Db 1040 ACTTGGCCCTTGTCTGCTCTAGTGTCTTCCCTTGAATGGGAAGAGTGTGGCCAA 1099

Qy 1206 AATCCCATCTTCTTGCACCTCAACGCTCTGTGGCTCAGGGCTGGGGTGGCAGAGGAGGC 1265

Db 1100 AATCCCATCTTCTTGCACCTCAACGCTCTGTGGCTCAGGGCTGGGGTGGCAGAGGAGGC 1159

Qy 1266 CTTACCTTATATCTGTGTGTGTTATCCAGGGCTCCAGACTTCTCTCTGTGCTGCCCCAC 1325

Db 1160 CTTACCTTATATCTGTGTGTGTTATCCAGGGCTCCAGACTTCTCTCTGTGCTGCCCCAC 1219

Qy 1326 TGCACCTCTCCCTTATATCTATCTCTCTCTGCTCCAGCTCCAGCTCCAGCTTGGCTTCTG 1385

Db 1220 TGCACCTCTCCCTTATATCTATCTCTCTCTGCTCCAGCTCCAGCTTGGCTTCTG 1279

Qy 1386 TCCCTCTCTGGGTCTATCCCTCACTCTGACTCTGACTATGACGAGCAAGAACACAGGGCC 1445

Db 1280 TCCCTCTCTGGGTCTATCCCTCACTCTGACTCTGACTATGACGAGCAAGAACACAGGGCC 1339

Qy 1446 TGGCCAGTGGATTTTCATGCTGATCATTTAAAAAGAAAAATCGAACCAAAAAA 1505

Db 1340 TGGCCAGTGGATTTTCATGCTGATCATTTAAAAAGAAAAATCGAACCAAAAAA 1399

RESULT 7

AC003042/c

LOCUS

DEFINITION

AC003042

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 102818)

Birren,B., Fasman,K., Linton,L., Nusbaum,C. and Lander,E.

Homo sapiens chromosome 17, clone HCIT75G16, complete sequence.

Unpublished

2 (bases 1 to 102818)

Birren,B., Fasman,K., McKernan,K., Munro,C., Nusbaum,C., Richardson,P., Lander,E., Baldwin,J., Barna,N., Cantu,C., Chang,A., Cooke,P., Daly,M.J., Devon,K., Dewar,K., DuRette,B., Forrest,C., Gage,D., Gensheimer,S., Geraghty,K., Gilman,T., Hagos,B., Halphen,I., Harris,K., Howland,J.C., Huang,J., Hui,L., Jacotot,L., Kirby,A., Lane,M., Mackenzie,J., Marquis,N., McDermott,J., Molla,M., Morrow,J., Nachman,A., Naylor,J., O'Connor,T., Olotu,A., Peterson,K., Roberts,D., Rollins,G., Sarnaik,A., Shiu,P., Shyam,R., Stilwell,J., Stone,C., Strickland,C., Sydney,K., Tang,L., Zemtseva,I. and Zody,M.

Direct Submission

Submitted (31-OCT-1997) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 102818)

Birren,B., Fasman,K., Linton,L., Nusbaum,C., Lander,E., Allen,N., Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Boatin,C., Boutwell,C., Brown,A., Castle,A., Cerny,J., Cooke,P., Depayre,E., Devon,K., Dewar,K., Donelan,L., Etemadi,S., Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D., Gardyna,S., Gensheimer,S., Geraghty,K., Gilman,T., Grant,G., Hagos,B., Harris,K., Horton,L., Howland,J.C., Hui,L., Jacotot,L., Kann,L., Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., Melcham,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J., Nachman,A., Naylor,J., Niloff,M., O'Connor,T., Pavlin,B., Peterson,K., Riley,R., Roberts,D., Rossello,R., Roy,A., Shyam,R., Stange-Thomann,N., Stilwell,J., Stojanovic,N., Stone,C.,

Strickland,C., Subramanian,A., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A., Wang,B., Wheeler,J., Wu,Y., Ye,W.J., Zhao,J. and Zody,M.

Direct Submission

Submitted (14-JUL-1998) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jul 14, 1998 this sequence version replaced gi:3294535.

All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html.

location/Qualifiers

1..102818

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="17"

/map="17"

/clones="HCIT75G16"

/clone_lib="Research Genetics/Cal Tech CITB978SK-B (plates 1-194)"

495..778

/rpt_family="AluJb"

complement(1357..1409)

/rpt_family="AT_rich"

1737..1800

/rpt_family=" (CA)n"

1911..1964

/rpt_family="MIR"

2074..2216

/rpt_family="MIR"

complement(2965..2985)

/rpt_family="AT_rich"

3225..3255

/rpt_family=" (GA)n"

3352..3446

/rpt_family=" (GGA)n"

complement(3681..3713)

/rpt_family="AT_rich"

complement(3966..4112)

/rpt_family="MIR"

complement(4970..5069)

/rpt_family="L2"

5648..5772

/rpt_family="MIR"

5957..5996

/rpt_family="MIR"

complement(6226..6270)

/rpt_family=" (TGG)n"

6890..6916

/rpt_family="GC_rich"

7990..8224

/rpt_family="AluSx"

8230..8737

/rpt_family="LIMB8"

8742..9020

/rpt_family="LIMB8"

complement(9027..9152)

/rpt_family="MER5A"

10394..10459

/rpt_family="L2"

complement(10478..10504)

/rpt_family=" (CAT)n"

complement(10582..10881)

/rpt_family="AluSg"

complement(10942..11013)

/rpt_family="MIR"

complement(11900..11931)

/rpt_family="MIR"

12398..12694

/rpt_family="AluJb"

complement(12919..13084)

/rpt_family="FRAM"

14235..14354

[illegible]

REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS	1 Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R., Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahari, K. and Masuho, Y.
TITLE	Full-length cDNA sequences
JOURNAL	Patent: EP 1347046-A 1353 24-SEP-2003; Research Association for Biotechnology (JP)
FEATURES	Location/Qualifiers
source	1. 2547 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"
ORIGIN	
Query Match	42.8%; Score 644; DB 6; Length 2547;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 644; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	850 CAGATTGGAGACATCCAGATGAGGCCCGGAGCAGGTGACCTAGTACTGTGGAGCTC 909
Db	1904 CAGATTGGAGACATCCAGATGAGGCCCGGAGCAGGTGACCTAGTACTGTGGAGCTC 1963
QY	910 CTCCTCCCTCCCGACCCCTTATGGCTTGCCTCGCTCGCTGGATTTAGTGTGATTT 969
Db	1964 CTCCTCCCTCCCGACCCCTTATGGCTTGCCTCGCTCGCTGGATTTAGTGTGATTT 2023
QY	970 CTGGATCAGGGATACCACTTCTCTCCACACCCCGACAGGGCTAGAAAATTTGTTTG 1029
Db	2024 CTGGATCAGGGATACCACTTCTCTCCACACCCCGACAGGGCTAGAAAATTTGTTTG 2083
QY	1030 AGATTTTATATCATCTTGTCAAAATTCAGTTGTAAATGTGAAAATGGCGTGGGA 1089
Db	2084 AGATTTTATATCATCTTGTCAAAATTCAGTTGTAAATGTGAAAATGGCGTGGGA 2143
QY	1090 AAGGAGGTGGTGTCCCTAATCTTTTACTTGTAACTTGTTCCTGTGCCCTGGGCATT 1149
Db	2144 AAGGAGGTGGTGTCCCTAATCTTTTACTTGTAACTTGTTCCTGTGCCCTGGGCATT 2203
QY	1150 GGCCCTTGTCTGCTCTCAGTGTCTTCCCTTGCATGGGAAAGGAGTGTGGCCAAATC 1209
Db	2204 GGCCCTTGTCTGCTCTCAGTGTCTTCCCTTGCATGGGAAAGGAGTGTGGCCAAATC 2263
QY	1210 CCCATCTTCTCAGCTCAACCTCAACCTGTGGCTCAGGGCTGGGTGGCAGAGGAGGCTTC 1269
Db	2264 CCCATCTTCTCAGCTCAACCTCAACCTGTGGCTCAGGGCTGGGTGGCAGAGGAGGCTTC 2323
QY	1150 GGCCCTTGTCTGCTCTCAGTGTCTTCCCTTGCATGGGAAAGGAGTGTGGCCAAATC 1209
Db	2204 GGCCCTTGTCTGCTCTCAGTGTCTTCCCTTGCATGGGAAAGGAGTGTGGCCAAATC 2263
QY	1210 CCCATCTTCTCAGCTCAACCTCAACCTGTGGCTCAGGGCTGGGTGGCAGAGGAGGCTTC 1269
Db	2264 CCCATCTTCTCAGCTCAACCTCAACCTGTGGCTCAGGGCTGGGTGGCAGAGGAGGCTTC 2323
ORIGIN	
Query Match	42.8%; Score 644; DB 9; Length 2547;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 644; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	850 CAGATTGGAGACATCCAGATGAGGCCCGGAGCAGGTGACCTAGTACTGTGGAGCTC 909
Db	1904 CAGATTGGAGACATCCAGATGAGGCCCGGAGCAGGTGACCTAGTACTGTGGAGCTC 1963
QY	910 CTCCTCCCTCCCGACCCCTTATGGCTTGCCTCGCTCGCTGGATTTAGTGTGATTT 969
Db	1964 CTCCTCCCTCCCGACCCCTTATGGCTTGCCTCGCTCGCTGGATTTAGTGTGATTT 2023
QY	970 CTGGATCAGGGATACCACTTCTCTCCACACCCCGACAGGGCTAGAAAATTTGTTTG 1029
Db	2024 CTGGATCAGGGATACCACTTCTCTCCACACCCCGACAGGGCTAGAAAATTTGTTTG 2083
QY	1030 AGATTTTATATCATCTTGTCAAAATTCAGTTGTAAATGTGAAAATGGCGTGGGA 1089
Db	2084 AGATTTTATATCATCTTGTCAAAATTCAGTTGTAAATGTGAAAATGGCGTGGGA 2143
QY	1090 AAGGAGGTGGTGTCCCTAATCTTTTACTTGTAACTTGTTCCTGTGCCCTGGGCATT 1149
Db	2144 AAGGAGGTGGTGTCCCTAATCTTTTACTTGTAACTTGTTCCTGTGCCCTGGGCATT 2203
QY	1150 GGCCCTTGTCTGCTCTCAGTGTCTTCCCTTGCATGGGAAAGGAGTGTGGCCAAATC 1209
Db	2204 GGCCCTTGTCTGCTCTCAGTGTCTTCCCTTGCATGGGAAAGGAGTGTGGCCAAATC 2263
QY	1210 CCCATCTTCTCAGCTCAACCTCAACCTGTGGCTCAGGGCTGGGTGGCAGAGGAGGCTTC 1269
Db	2264 CCCATCTTCTCAGCTCAACCTCAACCTGTGGCTCAGGGCTGGGTGGCAGAGGAGGCTTC 2323
QY	1270 ACCTTATATCTGTTGTTATCCAGGGCTCCAGACTTCTCTCTGCTGCTGCCCTGCACTGCA 1329
Db	2324 ACCTTATATCTGTTGTTATCCAGGGCTCCAGACTTCTCTCTGCTGCTGCCCTGCACTGCA 2383
QY	1330 CCTCTCTCCCTTATCTATCTCTCTCTCGGCTCCCGACGCCAGTCTTGTGCTTCTTGTCCC 1389
Db	2384 CCTCTCTCCCTTATCTATCTCTCTCTCGGCTCCCGACGCCAGTCTTGTGCTTCTTGTCCC 2443
QY	1390 CTCCTGGGGTATCTCTCTCACTCTGACTCTGACTATGGCAGCAGAACACACAGGGCTGGC 1449
Db	2444 CTCCTGGGGTATCTCTCTCACTCTGACTCTGACTATGGCAGCAGAACACACAGGGCTGGC 2503
QY	1450 CCAAGTGGATTTTATGTTGATCATTAAAAAAGAAAAATCGCAACC 1493
Db	2504 CCAAGTGGATTTTATGTTGATCATTAAAAAAGAAAAATCGCAACC 2547
RESULT 9	
LOCUS	AK096551
DEFINITION	Homo sapiens cDNA FLJ39232 fis, clone OCBF2007622.
ACCESSION	AK096551
VERSION	AK096551.1 GI:21756072
KEYWORDS	oligo capping; fis (full insert sequence).

```

QY 1270 ACCTTATATCTGTGTGTTATCCAGGGCTCCAGACTTCTCTCTGCTGCCCTCCACTGCA 1329
Db 2324 ACCTTATATCTGTGTGTTATCCAGGGCTCCAGACTTCTCTCTGCTGCCCTCCACTGCA 2383
QY 1330 CCCTTCCCTCCCTTATCTATCTCTCTCTGCTGCCCTCCAGGGCTCCACTGCTTCTGTGCC 1389
Db 2384 CCCTTCCCTCCCTTATCTATCTCTCTCTGCTGCCCTCCAGGGCTCCACTGCTTCTGTGCC 2443
QY 1390 CTCCTGGGCTATCCCTCCACTCTGACTCTGACTTGGCAGCAGACACCGAGGGCTGGC 1449
Db 2444 CTCCTGGGCTATCCCTCCACTCTGACTCTGACTTGGCAGCAGACACCGAGGGCTGGC 1449
QY 1450 CCAGTGGATTTTCATGGTGATCATTTAAAGAAAAAATGCAACC 1493
Db 2504 CCAGTGGATTTTCATGGTGATCATTTAAAGAAAAAATGCAACC 2547

RESULT 10
AC110594
LOCUS
DEFINITION Homo sapiens chromosome 17, clone CTD-2193J24, complete sequence.
ACCESSION AC110594
VERSION AC110594.5 GI:22830329
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 159490)
Birren,B., Nussbaum,C. and Lander,E.
Homo sapiens chromosome 17, clone CTD-2193J24
Unpublished
2 (bases 1 to 159490)
Birren,B., Linton,L., Nussbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferrelira,P., FitzHugh,W., Gage,D., Galaagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRoque,K., Lamazares,R.,
Landers,T., Lehoczyk,J., Levine,R., Liu,G., MacLean,C.,
Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (13-FEB-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 159490)
Birren,B., Nussbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cooke,P., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferrelira,P., FitzGerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P.,

```

Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.

TITLE JOURNAL AUTHORS

Submitted (17-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 159490)

Birren,B., Nussbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cooke,P., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferrelira,P., FitzGerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (13-SEP-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

On Sep 13, 2002 this sequence version replaced gi:22296745.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WTHR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L25381

Center clone name: 2193_J24

Only the first 159.5 kilobases of this clone are being submitted.

The remainder overlaps accession number AC003042 [WICGR project

L260].

FEATURES source

Location/Qualifiers

1. 159490
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="CTD-2193J24"
complement(5095..5617)
/rpt_family="L1MD3"
8205..8396
/rpt_family="AluSg/x"
8399..8482
/rpt_family="LTR65"
8505..8633
/rpt_family="FLAM_A"
complement(8643..8779)
/rpt_family="AluJo"
8780..8986
/rpt_family="LTR65"
complement(8988..9076)
/rpt_family="MIR"
complement(10090..10879)
/rpt_family="HAL1"
complement(10880..10989)
/rpt_family="FLAM_C"

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region	complement(10990..11042) /rpt family="HAL1"	Query Match	39.5%; Score 595; DB 9; Length 159490;
repeat_region	complement(11056..11357) /rpt family="AluSx"	Best Local Similarity	99.8%; Pred. No. 0;
repeat_region	complement(11367..11430) /rpt family="HAL1"	Matches	645; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
repeat_region	12628..13088 /rpt family="AluSx"	QY	850 CAGATTGGAGACATCCAGATGAGGCCACGAGCAGAGTACCTAGTACTGTGGAGCTC 909
repeat_region	15961..16045 /rpt family="GA-rich"	Db	150790 CAGATTGGAGACATCCAGATGAGGCCACGAGCAGAGTACCTAGTACTGTGGAGCTC 150849
repeat_region	16215..16446 /rpt family="AluJb"	QY	910 CTCCTTCCCTCCCAACCTTCATGGCTTGCCTCTGCTCTGGATTTTAGTGTGTTT 969
repeat_region	complement(16770..17108) /rpt family="AluJb"	Db	150850 CTCCTTCCCTCCCAACCTTCATGGCTTGCCTCTGCTCTGGATTTTAGTGTGTTT 150909
repeat_region	18252..18570 /rpt family="AluJc"	QY	970 CTGATCAGGGATACACATTCCTGTCCACACCCGACGAGGCTAGAAATTTGTTT 1029
repeat_region	complement(18770..18989) /rpt family="MIR"	Db	150910 CTGGATCAGAGTACACATTCCTGTCCACACCCGACGAGGCTAGAAATTTGTTT 150969
unsure	complement(18887..19024) /note="single clone coverage"	QY	1030 AGATTTTATATCATCTTGTCAAAATGCTTCAGTTGTAAATGTGAAAAATGGCTGGGA 1089
repeat_region	complement(19056..19189) /rpt family="LTR28"	Db	150970 AGATTTTATATCATCTTGTCAAAATGCTTCAGTTGTAAATGTGAAAAATGGCTGGGA 151029
unsure	complement(19236..19253) /note="<30 qual SNGL region"	QY	1090 AAGGAGTGGTGTCCCTAAATTTGTTTACTTGTAACTTGTCTTGTGCCCCCTGGGCAC 1149
repeat_region	complement(19244..20030) /rpt family="LTRLB"	Db	151030 AAGGAGTGGTGTCCCTAAATTTGTTTACTTGTAACTTGTCTTGTGCCCCCTGGGCAC 151089
unsure	19818..20633 /note="PCR product sequence only"	QY	1150 GGCCCTTGTCTGCTCTCAGTGTCTTCCCTTTGACATGGGAAAGGAGTTGTGGCCAAATC 1209
repeat_region	complement(20031..20388) /rpt family="HUERS-P2"	Db	151090 GGCCCTTGTCTGCTCTCAGTGTCTTCCCTTTGACATGGGAAAGGAGTTGTGGCCAAATC 151149
repeat_region	complement(20468..21174) /rpt family="HUERS-P2"	QY	1210 CCCATCTTTGACCTCAAGCTCTGTGGCTCAGGGCTGGGGTGGCAGAGGAGCCCTTC 1269
repeat_region	complement(21174..24069) /rpt family="HUERS-P2"	Db	151150 CCCATCTTTGACCTCAAGCTCTGTGGCTCAGGGCTGGGGTGGCAGAGGAGCCCTTC 151209
repeat_region	24070..24364 /rpt family="AluSg"	QY	1270 ACCTTATATCTGTGTGTTTATCCAGGGCTCCAGACTTCTCTCTGCTGCCCTGCA 1329
repeat_region	complement(24365..25300) /rpt family="HUERS-P2"	Db	151210 ACCTTATATCTGTGTGTTTATCCAGGGCTCCAGACTTCTCTCTGCTGCCCTGCA 151269
repeat_region	complement(25436..25569) /rpt family="LTR28"	QY	1330 CCCTCTCCCTTATCTATCTCTCGGCTCCCGAGCCAGTCTTGGCTCTTGTGCTCC 1389
repeat_region	complement(25594..26436) /rpt family="LTRLB"	Db	151270 CCCTCTCCCTTATCTATCTCTCGGCTCCCGAGCCAGTCTTGGCTCTTGTGCTCC 151329
repeat_region	complement(26936..27052) /rpt family="LIME1"	QY	1390 CTCCTGGGGTCATCCCTCCACTCTGACTCTGACTATGGCAGCAGAACACAGGGCTGGC 1449
repeat_region	complement(27053..27348) /rpt family="AluSc"	Db	151330 CTCCTGGGGTCATCCCTCCACTCTGACTCTGACTATGGCAGCAGAACACAGGGCTGGC 151389
repeat_region	complement(27349..28233) /rpt family="LIME1"	QY	1450 CCAGTGGATTTTCATGGTGATCATTTAAAAAGAAAAATGCAACCAA 1495
repeat_region	complement(28259..28436) /rpt family="AluJo"	Db	151390 CCAGTGGATTTTCATGGTGATCATTTAAAAAGAAAAATGCAACCAA 151435
repeat_region	complement(28442..28555) /rpt family="LIME1"	RESULT 11	
repeat_region	28556..28643 /rpt family="FLAM A"	AX768992	AX768992 2077 bp DNA linear PAT 02-JUL-2003
repeat_region	complement(28644..28949) /rpt family="AluSx"	LOCUS	Sequence 109 from Patent WO02098917.
repeat_region	28950..28994 /rpt family="FLAM A"	DEFINITION	AX768992
repeat_region	complement(29006..29119) /rpt family="LIME1"	ACCESSION	AX768992
repeat_region	complement(29120..29418) /rpt family="AluY"	VERSION	AX768992.1 GI:32437160
repeat_region	complement(29419..29571) /rpt family="LIME1"	KEYWORDS	Homo sapiens (human)
repeat_region	complement(29583..29744) /rpt family="LIME1"	SOURCE	Homo sapiens
repeat_region	complement(29794..29997) /rpt family="MER20"	ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
repeat_region	complement(30024..30125) /rpt family="LIME1"	REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
repeat_region	complement(30126..30263) /rpt family="FLAM A"	AUTHORS	1 Guo X., Fernandes E., Li L., Kekuda R., Liu Y., Leite M., Spytek K.A., Ji W., Casman S.J., Boldog F.L., Patturajan M., Vernet C.A., Ballinger R.A., Malyankar U.M., Tchernev V.T., Blalock A.D., Gusev V.Y., Rastelli L., Mezes P.D., Ellerman K., Heyes M., Herrmann J.L., Shimkets R.A., Ioline N., Pena C.E., Shenoy S.G., Taupier R.J., Gerlach V. and Gorman L.
repeat_region	complement(30264..30551)	TITLE	Human proteins and nucleic acids encoding same
		JOURNAL	Patent: WO 02098917-A 109 12-DEC-2002;
		FEATURES	Curagen Corporation (US)
		source	1..2077

/organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"			
ORIGIN	Query Match Best Local Similarity 37.8%; Score 569; DB 6; Length 2077; Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	127 ATGGAGCGGTGGCGGACCGGCTGGCGCTGGTGACGCGGGGCTCGGGGGGCATCGGGCG 186		
Db	1 ATGGAGCGGTGGCGGACCGGCTGGCGCTGGTGACGCGGGGCTCGGGGGGCATCGGGCG 60		
QY	187 GCCGTGGCCCGGCGCTGGTCCAGCAGGAGCTGAAGAGTGGTGGGCTGGCCCGCAGCTGTG 246		
Db	61 GCCGTGGCCCGGCGCGCTGGTCCAGCAGGAGCTGAAGAGTGGTGGGCTGGCCCGCAGCTGTG 120		
QY	247 GGCAACATCGAGGAGCTGGCTGCTGAATGTAAGAGTGACGGCTACCCGGGACTTTGATC 306		
Db	121 GGCAACATCGAGGAGCTGGCTGCTGAATGTAAGAGTGACGGCTACCCGGGACTTTGATC 180		
QY	307 CCCTACAGATGTGACCTATCAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGT 366		
Db	181 CCCTACAGATGTGACCTATCAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGT 240		
QY	367 TCTCAGCAGCGGTGTAGACATCTGCATCAACAATGCTGGCTTGGCCCGCCTGACACC 426		
Db	241 TCTCAGCAGCAGCGGTGTAGACATCTGCATCAACAATGCTGGCTTGGCCCGCCTGACACC 300		
QY	427 CTGCTCTCAGGAGCAGCAGTGGTTGGAAGGACATGTTCAATGTGAACGTGCTGGCCCTC 486		
Db	301 CTGCTCTCAGGAGCAGCAGTGGTTGGAAGGACATGTTCAATGTGAACGTGCTGGCCCTC 360		
QY	487 AGCATCTGCACACGGGAAGCTACCACTCCATGAAGAGCGGAATGTGGACGATGGGCAC 546		
Db	361 AGCATCTGCACACGGGAAGCTACCACTCCATGAAGAGCGGAATGTGGACGATGGGCAC 420		
QY	547 ATCATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGTCTGACCCACTTC 606		
Db	421 ATCATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGTCTGACCCACTTC 480		
QY	607 TATAGTGCACCAAGTATGCGGTCACTGCGTGCACAGAGGACCTGAGGCAAGAGCTTCGG 666		
Db	481 TATAGTGCACCAAGTATGCGGTCACTGCGTGCACAGAGGACCTGAGGCAAGAGCTTCGG 540		
QY	667 GAGGCCAGACCCACATCCGAGCCACGTG 695		
Db	541 GAGGCCAGACCCACATCCGAGCCACGTG 569		
RESULT 12			
AR184145			
LOCUS	AR184145 569 bp DNA linear PAT 20-APR-2002		
DEFINITION	Sequence 89 from patent US 6342581.		
ACCESSION	AR184145		
VERSION	AR184145.1 GI:20228114		
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 569)		
AUTHORS	Rosen, C.A., Ruben, S.M., Olsen, H.S. and Ebner, R.		
TITLE	Secreted protein HLFp03		
JOURNAL	Patent: US 6342581-A 89 29-JAN-2002;		
FEATURES	Location/Qualifiers		
source	1..569		
/organism="unknown"			
/mol_type="unassigned DNA"			
ORIGIN	Query Match Best Local Similarity 29.6%; Score 445; DB 6; Length 569; Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		

QY	998	ACACCCGACCGAGGGCTAGAAAAATTTGTTTGTGAGATTTTTATATCATCTTGTCAAATTC	1057
Db	42	ACACCCGACCGAGGGCTAGAAAAATTTGTTTGTGAGATTTTTATATCATCTTGTCAAATTC	101
QY	1058	TTCAAGTTGTAAATGTGAAAAATGGGCTGGGAAAGAGGAGTGGTCCCTAAATGTTTTTAC	1117
Db	102	TTCAAGTTGTAAATGTGAAAAATGGGCTGGGAAAGAGGAGTGGTCCCTAAATGTTTTTAC	161
QY	1118	TTGTTAACTTGTCTTGTGTGCCCTGGGCACCTTGGCCCTTGTCTGCTCTCAGTGTCTTCCC	1177
Db	162	TTGTTAACTTGTCTTGTGTGCCCTGGGCACCTTGGCCCTTGTCTGCTCTCAGTGTCTTCCC	221
QY	1178	TTTGACATGGGAAAGAGTGTGSCCAAATAATCCCATCTTCTTGTGCACTCAACGCTCTGTG	1237
Db	222	TTTGACATGGGAAAGAGTGTGSCCAAATAATCCCATCTTCTTGTGCACTCAACGCTCTGTG	281
QY	1238	GCTCAGGGCTGGGGTGGCAGAGGAGGCGCTTACACCTTATATCTGTGTGTTATCCAGGGC	1297
Db	282	GCTCAGGGCTGGGGTGGCAGAGGAGGCGCTTACACCTTATATCTGTGTGTTATCCAGGGC	341
QY	1298	TCCAGACTTCTCTCTCTGCTGCCCACTGCACCCCTCTCCCTTATCTATCTCTCTCTC	1357
Db	342	TCCAGACTTCTCTCTCTGCTGCCCACTGCACCCCTCTCCCTTATCTATCTCTCTCTC	401
QY	1358	GGCTCCCCAGCCAGTCTTGGCTTCTTGTCCCTCTCGGGGTATCCCTCAGCTCTGACT	1417
Db	402	GGCTCCCCAGCCAGTCTTGGCTTCTTGTCCCTCTCGGGGTATCCCTCAGCTCTGACT	461
QY	1418	CTGACTATGGCAGCAGAACACACAGG 1442	
Db	462	CTGACTATGGCAGCAGAACACACAGG 486	
RESULT 13			
BD192674			
LOCUS	BD192674 569 bp DNA linear PAT 17-JUL-2003		
DEFINITION	123 human secreted proteins.		
ACCESSION	BD192674		
VERSION	BD192674.1 GI:33002413		
KEYWORDS	JP 2002513295-A/88.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 569) Fischer, C.L., Rosen, C.A., Soppet, D.R., Ruben, S.M., Kyaw, H., Li, Y., Zeng, Z., Lafleur, D.W., Moore, P.A., Shi, Y., Ols, H.S., Ebner, R. and Brewer, L.A.		
TITLE	123 human secreted proteins		
JOURNAL	Patent: JP 2002513295-A 88 08-MAY-2002;		
COMMENT	HUMAN GENOME SCIENCES INC OS Homo sapiens (human) PN JP 2002513295-A/88 PD 08-MAY-2002 PF 07-JUL-1998 JP 1999508744 PR 08-JUL-1997 US 60/051926, 08-JUL-1997 US 60/051929 PR 08-JUL-1997 US 60/052793, 08-JUL-1997 US 60/051925 PR 08-JUL-1997 US 60/051931, 08-JUL-1997 US 60/051932 PR 08-JUL-1997 US 60/052803, 08-JUL-1997 US 60/052732 PR 08-JUL-1997 US 60/051916, 08-JUL-1997 US 60/051930 PR 08-JUL-1997 US 60/051918, 08-JUL-1997 US 60/051920 PR 08-JUL-1997 US 60/051919, 08-JUL-1997 US 60/051928 PR 08-JUL-1997 US 60/052795, 08-JUL-1997 US 60/052733 PR 18-AUG-1997 US 60/055948, 18-AUG-1997 US 60/055722 PR 18-AUG-1997 US 60/055723, 18-AUG-1997 US 60/055949 PR 18-AUG-1997 US 60/055964, 18-AUG-1997 US 60/055947 PR 18-AUG-1997 US 60/055950, 18-AUG-1997 US 60/055953 PR 18-AUG-1997 US 60/056360, 18-AUG-1997 US 60/055684 PR 18-AUG-1997 US 60/055984, 18-AUG-1997 US 60/055954 PR 12-SEP-1997 US 60/058785, 12-SEP-1997 US 60/058664 PR 12-SEP-1997 US 60/058660, 12-SEP-1997 US 60/058661 PI L FISCHER, CRAIG A ROSEN, DANIEL R SOPPET, STEVEN M RUBEN, PI HUA		

JOURNAL

COMMENT

Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jul 13, 2000 this sequence version replaced gi:6446748.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L855
 Center clone name: 2047_D_24

* NOTE: This record contains 177 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

* 1 841: contig of 841 bp in length
 * 842 941: gap of 100 bp
 * 942 1778: contig of 837 bp in length
 * 1779 1878: gap of 100 bp
 * 1879 2729: contig of 851 bp in length
 * 2730 2829: gap of 100 bp
 * 2830 3681: contig of 852 bp in length
 * 3682 3781: gap of 100 bp
 * 3782 4620: contig of 839 bp in length
 * 4621 4720: gap of 100 bp
 * 4721 5544: contig of 824 bp in length
 * 5545 5644: gap of 100 bp
 * 5645 6505: contig of 861 bp in length
 * 6506 6605: gap of 100 bp
 * 6606 7451: contig of 846 bp in length
 * 7452 7551: gap of 100 bp
 * 7552 8385: contig of 834 bp in length
 * 8386 8485: gap of 100 bp
 * 8486 9315: contig of 830 bp in length
 * 9316 9415: gap of 100 bp
 * 9416 10262: contig of 847 bp in length
 * 10263 10362: gap of 100 bp
 * 10363 11228: contig of 866 bp in length
 * 11229 11328: gap of 100 bp
 * 11329 12180: contig of 852 bp in length
 * 12181 12280: gap of 100 bp
 * 12281 13121: contig of 841 bp in length
 * 13122 13221: gap of 100 bp
 * 13222 14070: contig of 849 bp in length
 * 14071 14170: gap of 100 bp
 * 14171 15030: contig of 860 bp in length
 * 15031 15130: gap of 100 bp
 * 15131 15974: contig of 844 bp in length
 * 15975 16074: gap of 100 bp
 * 16075 16900: contig of 826 bp in length
 * 16901 17000: gap of 100 bp
 * 17001 17835: contig of 835 bp in length
 * 17836 17935: gap of 100 bp
 * 17936 18775: contig of 840 bp in length
 * 18776 18875: gap of 100 bp
 * 18876 19721: contig of 846 bp in length
 * 19722 19821: gap of 100 bp
 * 19822 20666: contig of 845 bp in length
 * 20667 20766: gap of 100 bp
 * 20767 21625: contig of 859 bp in length
 * 21626 21725: gap of 100 bp
 * 21726 22617: contig of 892 bp in length

22618 22717: gap of 100 bp
 22718 23551: contig of 834 bp in length
 23552 23651: gap of 100 bp
 23652 24492: contig of 841 bp in length
 24493 24592: gap of 100 bp
 24593 25450: contig of 858 bp in length
 25451 25550: gap of 100 bp
 25551 26411: contig of 861 bp in length
 26412 26511: gap of 100 bp
 26512 27357: contig of 846 bp in length
 27358 27457: gap of 100 bp
 27458 28297: contig of 840 bp in length
 28298 28397: gap of 100 bp
 28398 29223: contig of 826 bp in length
 29224 29323: gap of 100 bp
 29324 30171: contig of 848 bp in length
 30172 30271: gap of 100 bp
 30272 31100: contig of 829 bp in length
 31101 31200: gap of 100 bp
 31201 32044: contig of 844 bp in length
 32045 32144: gap of 100 bp
 32145 32989: contig of 845 bp in length
 32990 33089: gap of 100 bp
 33090 33951: contig of 862 bp in length
 33952 34051: gap of 100 bp
 34052 34908: contig of 857 bp in length
 34909 35009: gap of 100 bp
 35009 35851: contig of 843 bp in length
 35852 35951: gap of 100 bp
 35952 36798: contig of 847 bp in length
 36799 36898: gap of 100 bp
 36899 37756: contig of 858 bp in length
 37757 37856: gap of 100 bp
 37857 38702: contig of 846 bp in length
 38703 38802: gap of 100 bp
 38803 39635: contig of 833 bp in length
 39636 39735: gap of 100 bp
 39736 40577: contig of 842 bp in length
 40578 40677: gap of 100 bp
 40678 41508: contig of 831 bp in length
 41509 42452: contig of 844 bp in length
 42453 42552: gap of 100 bp
 42553 43405: contig of 853 bp in length
 43406 43505: gap of 100 bp
 43506 44360: contig of 855 bp in length
 44361 44460: gap of 100 bp
 44461 45294: contig of 834 bp in length
 45295 45394: gap of 100 bp
 45395 46256: contig of 862 bp in length
 46257 46356: gap of 100 bp
 46357 47221: contig of 865 bp in length
 47222 47321: gap of 100 bp
 47322 48156: contig of 835 bp in length
 48157 48256: gap of 100 bp
 48257 49090: contig of 834 bp in length
 49091 49190: gap of 100 bp
 49191 50043: contig of 853 bp in length
 50044 50143: gap of 100 bp
 50144 50989: contig of 846 bp in length
 50990 51089: gap of 100 bp
 51089 51925: contig of 836 bp in length
 51926 52025: gap of 100 bp
 52026 52874: contig of 849 bp in length
 52875 52974: gap of 100 bp
 52975 53824: contig of 850 bp in length
 53825 53924: gap of 100 bp
 53925 54887: contig of 863 bp in length
 54888 55740: contig of 853 bp in length
 55741 55840: gap of 100 bp
 55841 56692: contig of 852 bp in length
 56693 56792: gap of 100 bp

		57654: contig of 862 bp in length	
*	56793	57654: gap of 100 bp	
*	57655	58595: contig of 841 bp in length	
*	57755	58696: gap of 100 bp	
*	58696	59533: contig of 838 bp in length	
*	59534	59633: gap of 100 bp	
*	59634	60479: contig of 846 bp in length	
*	60480	60579: gap of 100 bp	
*	60580	61430: contig of 851 bp in length	
*	61431	61530: gap of 100 bp	
*	61531	62371: contig of 841 bp in length	
*	62372	62471: gap of 100 bp	
*	62472	63313: contig of 842 bp in length	
*	63314	63413: gap of 100 bp	
*	63414	64260: contig of 847 bp in length	
*	64261	64360: gap of 100 bp	
*	64361	65225: contig of 865 bp in length	
*	65226	65325: gap of 100 bp	
*	65326	66181: contig of 856 bp in length	
*	66182	66281: gap of 100 bp	
*	66282	67127: contig of 846 bp in length	
*	67128	67227: gap of 100 bp	
*	67228	68088: contig of 861 bp in length	
*	68089	68188: gap of 100 bp	
Query Match 15.1%; Score 227; DB 2; Length 169713;			
Best Local Similarity 99.7%; Pred. No. 1.8e-117;			
Matches 347; Conservative 0; Mismatches 0; Indels 1; Gaps 1;			
QY	966	ATTCTGGATCAGGGATACCACTTCCTGTCCACACCCCGACCGGGGCTAGAAAAATTG	1025
Db	45931		45872
QY	1026	TTTGAGATTTTATATCATCTTTGTCAAAATGCTTCAGTTGTAAATGTGAAAAATGGGCTG	1085
Db	45871		45812
QY	1086	GGGAAAGAGGTGGTGTCCCTAAATGTTTACTTGTAACTTGTTCCT-TGTCCCTGGG	1144
Db	45811	GGGAAAGAGGTGGTGTCCCTAAATGTTTACTTGTAACTTGTTCCTATGTATGCTCCCTGGG	45752
QY	1145	CACCTGGCTTTGTCTGTCTCAGTGTCTTCCTTTGACATGGGAAAGAGTTGTGGCCA	1204
Db	45751	CACCTGGCTTTGTCTGTCTCAGTGTCTTCCTTTGACATGGGAAAGAGTTGTGGCCA	45692
QY	1205	AAATCCCATCTTTGTGCACCTCAACGTCGTGGCTCAGGGTGGGTCAGAGGAGG	1264
Db	45691	AAATCCCATCTTTGTGCACCTCAACGTCGTGGCTCAGGGTGGGTCAGAGGAGG	45632
QY	1265	CCTTCACCTTATATCTGTGTTCTTATCCAGGGCTCCAGACTTCCTCCT	1312
Db	45631	CCTTCACCTTATATCTGTGTTCTTATCCAGGGCTCCAGACTTCCTCCT	45584

Search completed: September 16, 2004, 16:23:38
Job time : 5841 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2004, 12:18:50 ; Search time 3656 Seconds
(without alignments)
12292.829 Million cell updates/sec

Title: US-10-063-735-127
Perfect score: 1505
Sequence: 1 cgcggatcgaccacgaagcag.....tcgcaacccccaaaaa 1505

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 10

Total number of hits satisfying chosen parameters: 32528767

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

EST:

- 1: em_estba:**
- 2: em_esthum:**
- 3: em_estin:**
- 4: em_estmu:**
- 5: em_estov:**
- 6: em_estpl:**
- 7: em_estro:**
- 8: em_hc:**
- 9: gb_est1:**
- 10: gb_est2:**
- 11: gb_est3:**
- 12: gb_est4:**
- 13: gb_est5:**
- 14: gb_est6:**
- 15: em_estfun:**
- 16: em_estom:**
- 17: em_gss_hum:**
- 18: em_gss_inv:**
- 19: em_gss_pln:**
- 20: em_gss_vrt:**
- 21: em_gss_fun:**
- 22: em_gss_mam:**
- 23: em_gss_mus:**
- 24: em_gss_pro:**
- 25: em_gss_rod:**
- 26: em_gss_phg:**
- 27: em_gss_vrl:**
- 28: gb_gss1:**
- 29: gb_gss2:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	675	44.9	829	12	BG324477 602422421
2	668	44.4	1010	12	BM471232
3	659	43.8	679	10	BF979853 602288368
4	654	43.5	1071	13	BQ278156

78	263	17.5	376	14	T49180	T49180	yB09h06.x1
C 79	254	16.9	486	9	RA908673	AA908673	o104d06.s
C 80	254	16.9	535	14	W58347	W58347	zd25b06.b1
C 81	253	16.8	413	9	AA575928	AA575928	nm56b12.s
C 82	252	16.7	312	13	BA480776	BA480776	DKF2p686C
C 83	249	16.5	352	10	BF375427	BF375427	MR4-ST024
C 84	248	16.5	374	10	AG392371	AG392371	MR4-ST024
C 85	249	16.5	539	12	BG488598	BG488598	602534511
C 86	246	16.3	299	10	BA351512	BA351512	IL2-CT003
C 87	246	16.3	641	13	BA488597	BA488597	DKF2p686L
C 88	246	16.3	748	12	BI761329	BI761329	603044093
C 89	237	15.7	417	10	AW371786	AW371786	CM2-BT030
C 90	234	15.5	242	9	AI699226	AI699226	tx62e11.x
C 91	234	15.5	1171	12	BM800956	BM800956	AGECNCURT
C 92	228	15.2	511	9	AI356759	AI356759	qy22f10.x
C 93	228	15.1	940	12	BI818880	BI818880	603037354
C 94	220	14.6	651	9	AV712795	AV712795	AV712795
C 95	214	14.2	348	9	AA723851	AA723851	ah63306.s
C 96	214	14.2	395	10	AW371791	AW371791	CM2-BT030
C 97	210	14.0	333	14	D60924	D60924	HUM139B12B
C 98	205	13.6	567	14	W58472	W58472	zd25b06.y1
C 99	204	13.4	536	14	W58459	W58459	zd25b06.b1
C 100	201	13.4	465	12	BM825434	BM825434	K-EST0097

ALIGNMENTS

RESULT 1	
BG324477	829 bp mRNA linear EST 27-FEB-2001
LOCUS	G024242421F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4560485 5' ,
DEFINITION	mRNA sequence.
ACCESSION	BG324477
VERSION	BG324477.1 GI:13130914
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE	1 (bases 1 to 829)
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs@mail.nih.gov Tissue Procurement: DCTD/BTP cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLCMI269 row: j column: 06 High quality sequence stop: 755.

FEATURES	SOURCE
1. <i>General</i>	
2. <i>Specific</i>	
3. <i>Other</i>	
4. <i>Other</i>	
5. <i>Other</i>	
6. <i>Other</i>	
7. <i>Other</i>	
8. <i>Other</i>	
9. <i>Other</i>	
10. <i>Other</i>	
11. <i>Other</i>	
12. <i>Other</i>	
13. <i>Other</i>	
14. <i>Other</i>	
15. <i>Other</i>	
16. <i>Other</i>	
17. <i>Other</i>	
18. <i>Other</i>	
19. <i>Other</i>	
20. <i>Other</i>	
21. <i>Other</i>	
22. <i>Other</i>	
23. <i>Other</i>	
24. <i>Other</i>	
25. <i>Other</i>	
26. <i>Other</i>	
27. <i>Other</i>	
28. <i>Other</i>	
29. <i>Other</i>	
30. <i>Other</i>	
31. <i>Other</i>	
32. <i>Other</i>	
33. <i>Other</i>	
34. <i>Other</i>	
35. <i>Other</i>	
36. <i>Other</i>	
37. <i>Other</i>	
38. <i>Other</i>	
39. <i>Other</i>	
40. <i>Other</i>	
41. <i>Other</i>	
42. <i>Other</i>	
43. <i>Other</i>	
44. <i>Other</i>	
45. <i>Other</i>	
46. <i>Other</i>	
47. <i>Other</i>	
48. <i>Other</i>	
49. <i>Other</i>	
50. <i>Other</i>	
51. <i>Other</i>	
52. <i>Other</i>	
53. <i>Other</i>	
54. <i>Other</i>	
55. <i>Other</i>	
56. <i>Other</i>	
57. <i>Other</i>	
58. <i>Other</i>	
59. <i>Other</i>	
60. <i>Other</i>	
61. <i>Other</i>	
62. <i>Other</i>	
63. <i>Other</i>	
64. <i>Other</i>	
65. <i>Other</i>	
66. <i>Other</i>	
67. <i>Other</i>	
68. <i>Other</i>	
69. <i>Other</i>	
70. <i>Other</i>	
71. <i>Other</i>	
72. <i>Other</i>	
73. <i>Other</i>	
74. <i>Other</i>	
75. <i>Other</i>	
76. <i>Other</i>	
77. <i>Other</i>	
78. <i>Other</i>	
79. <i>Other</i>	
80. <i>Other</i>	
81. <i>Other</i>	
82. <i>Other</i>	
83. <i>Other</i>	
84. <i>Other</i>	
85. <i>Other</i>	
86. <i>Other</i>	
87. <i>Other</i>	
88. <i>Other</i>	
89. <i>Other</i>	
90. <i>Other</i>	
91. <i>Other</i>	
92. <i>Other</i>	
93. <i>Other</i>	
94. <i>Other</i>	
95. <i>Other</i>	
96. <i>Other</i>	
97. <i>Other</i>	
98. <i>Other</i>	
99. <i>Other</i>	
100. <i>Other</i>	

RESULT 2

BM471232	BM471232	1010 bp	mRNA	linear	EST 05-FEB-2002							
LOCUS	ACENECOURT_6478395	NIH_MGC_72	Homo sapiens	cdna clone IMAGE:5563124								
DEFINITION	5', mRNA sequence.											
ACCESSION	BM471232											
VERSION	BM471232.1	GI:18520274										
KEYWORDS	EST.											
SOURCE	Homo sapiens	Homo sapiens (human)										
ORGANISM	Homo sapiens											
REFERENCE	1 (bases 1 to 1010)											
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .											
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)											
JOURNAL	Unpublished (1999)											
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC/DCTDP/DTP											

Query Match	44.9%;	Score 675;	DB 12;	Length 829;
Best Local Similarity	100.0%;	Pred. No. 4.1e-219;		
Matches 675;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
45	GC	CGGGCTCAGCTCCTCGACCCCGCTGCTCGGGCTAGTCCAGCGAGCGGCGGACGGCGGCGG	104	
1	GC	CGGGCTCAGCTCCTCGACCCCGCTGCTCGGGCTAGTCCAGCGAGCGGCGGACGGCGGCGG	60	
105	TG	GGCCCATGGCCAGGCGCGGCATGAGCGGCTGGCGCGACCGGCTGGCGCTGGTGAACGGG	164	
61	TG	GGCCCATGGCCAGGCGCGGCATGAGCGGCTGGCGCGACCGGCTGGCGCTGGTGAACGGG	120	
165	GG	CTCGGGGGGATCGGGCGGGCGTGGCCCGGGCCCTGTTCAGCAGGAGACTGAAGGT	224	
121	GG	CTCGGGGGGATCGGGCGGGCGTGGCCCGGGCCCTGTTCAGCAGGAGACTGAAGGT	180	
225	GG	TGGGCTCGGCGCGGCACCTGTGGGCAACATCGAGGAGCTGGCTGCTGAATGTAAGAGTGC	284	
181	GG	TGGGCTCGGCGCGGCACCTGTGGGCAACATCGAGGAGCTGGCTGCTGAATGTAAGAGTGC	240	
285	AG	GCTACCCCGGGACTTTTGATCCCTACAGATGTGACCTATCAAATGAAGAGGACATCCT	344	
241	AG	GCTACCCCGGGACTTTTGATCCCTACAGATGTGACCTATCAAATGAAGAGGACATCCT	300	
345	CT	CCATGTTCTCAGCTATCCGTTCTCAGCAGCGGTGAGACATCTGCATCAACAATGC	404	
301	CT	CCATGTTCTCAGCTATCCGTTCTCAGCAGCGGTGAGACATCTGCATCAACAATGC	360	
405	TG	GCTTGGCGCGGCTGACACCCCTGCTCTCAGCGAGCACCAGTGGTTGGAAGGACATGTT	464	
361	TG	GCTTGGCGCGGCTGACACCCCTGCTCTCAGCGAGCACCAGTGGTTGGAAGGACATGTT	420	
465	CA	ATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCTTACCAGTCCATGAAGGA	524	
421	CA	ATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCTTACCAGTCCATGAAGGA	480	
525	GC	GAATGTGACGATGGGCACATCTTAACATCAATAGCATGTCTGGCCACCGAGTGT	584	
481	GC	GAATGTGACGATGGGCACATCTTAACATCAATAGCATGTCTGGCCACCGAGTGT	540	
585	AC	CCCTGCTGTGACCCACTTCTATAGTGGCCACCAAGTATGCCGTCTGCCTGCACAGA	644	
541	AC	CCCTGCTGTGACCCACTTCTATAGTGGCCACCAAGTATGCCGTCTGCCTGCACAGA	600	
645	GG	GACTTGAGCAAGAGCTTCGGGAGGCCCCAGACCCACATCCGAGGCCAGTGATCTCTCC	704	
601	GG	GACTTGAGCAAGAGCTTCGGGAGGCCCCAGACCCACATCCGAGGCCAGTGATCTCTCC	660	
705	AG	TGTGGTGGAGAC	719	
661	AG	TGTGGTGGAGAC	675	

ORIGIN

cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM12293 row: j column: 21
 High quality sequence stop: 561.
 Location/Qualifiers
 1..1010
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5563124"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_72"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
 Average insert size 2 kb. Library constructed by Life
 Technologies."

ORIGIN

Query Match 44.4%; Score 668; DB 12; Length 1010;
 Best Local Similarity 100.0%; Pred. No. 8.9e-217;
 Matches 668; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GAGCGCGCGGCTCAGCTCTCGACCCCGCTGTCTCGGCTAGTCAGGAGCGGACGGG 99
 DB 15 GAGCGCGCGGCGTCTCTCGACCCCGCTGTCTCGGCTAGTCAGGAGCGGACGGG 74
 QY 100 CGGCGTGGCGCCATGCGCGAGCGCGGATGAGCGGTGGCGCGACCGGCTGGCTGTG 159
 DB 75 CGGCGTGGCGCCATGCGCGAGCGCGGATGAGCGGTGGCGCGACCGGCTGGCTGTG 134
 QY 160 ACGGGGCGCTCGGGGGGATCGGCGCGCGCTGTGCGCGGCTGTGTCAGCAGGAGCTG 219
 DB 135 ACGGGGCGCTCGGGGGGATCGGCGCGCGCTGTGCGCGGCTGTGTCAGCAGGAGCTG 194
 QY 220 AAGTGTGTGGCTCGCGCGGCTGTGCGCGGATCGGCGGAGCTGTGCTGAATGTAAG 279
 DB 195 AAGTGTGTGGCTCGCGCGGCTGTGCGCGGATCGGCGGAGCTGTGCTGAATGTAAG 254
 QY 280 AGTCAGGCTACCCCGGAGCTTTGATCCCTACAGATGTGACCTTCAAAATGAAGAGAC 339
 DB 255 AGTCAGGCTACCCCGGAGCTTTGATCCCTACAGATGTGACCTTCAAAATGAAGAGAC 314
 QY 340 ATCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTGTAGACATCTGCATCAAC 399
 DB 315 ATCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTGTAGACATCTGCATCAAC 374
 QY 400 AATGCTGGCTTGGCGCGGCTGACACCTGCTCTCAGGCGACCGAGTGGTGAAGGAC 459
 DB 375 AATGCTGGCTTGGCGCGGCTGACACCTGCTCTCAGGCGACCGAGTGGTGAAGGAC 434
 QY 460 ATGTTCAATGTGAACGTGTGGCGCTCAGCATCTGCAACGGGAGCTTACAGTCCATG 519
 DB 435 ATGTTCAATGTGAACGTGTGGCGCTCAGCATCTGCAACGGGAGCTTACAGTCCATG 494
 QY 520 AAGAGCGGAATGTGACGATGGGCAATCATTAACATCAATAGCATGTCTGGCCACCGA 579
 DB 495 AAGAGCGGAATGTGACGATGGGCAATCATTAACATCAATAGCATGTCTGGCCACCGA 554
 QY 580 GTGTTACCCCTGTCTGTGACCCACTTCTATAGTGCACCAAGTATGCGGTCACTGCGCTG 639
 DB 555 GTGTTACCCCTGTCTGTGACCCACTTCTATAGTGCACCAAGTATGCGGTCACTGCGCTG 614
 QY 640 ACAGAGGAGCTGAGGCAAGAGCTTCGGGAGGCCAGACCCACATCCGAGCCAGCTGCATC 699
 DB 615 ACAGAGGAGCTGAGGCAAGAGCTTCGGGAGGCCAGACCCACATCCGAGCCAGCTGCATC 674
 QY 700 TCTCCAGG 707
 |||||

Db

675 TCTCCAGG 682

RESULT 3

BF979853/c

LOCUS

DEFINITION

BF979853

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BF979853

6022883681 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4374184 3',

mRNA sequence.

EST.

BF979853.1 GI:12347068

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 679)

NIH-MGC http://mgi.nci.nih.gov/.

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs@mail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki

Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM10037 row: g column: 17

High quality sequence start: 24

High quality sequence stop: 679.

Location/Qualifiers

1..679

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4374184"

/lab_host="DH10B"

/clone_lib="NIH_MGC_97"

/note="Organ: testis; Vector: pBluescriptR (modified

pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI

(gtcgag); Oligo-dT primed using primer

5'-TTTTTTTTTTTTTTVN-3', size-selected for average

insert size 2.2 kb and normalized to 5x. This is a

primary library enriched for full-length clones and

constructed using the Cap-trapper method (Carninci, in

preparation). Library constructed by M. Brownstein

(NIH/NHGRI, National Institutes of Health). Note: this is

a NIH_MGC Library."

ORIGIN

Query Match 43.8%; Score 659; DB 10; Length 679;
 Best Local Similarity 100.0%; Pred. No. 1.2e-213;
 Matches 659; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 767 CACCTATGAGCAATGAAGTGTCTCAACCCGAGGATGTGGCGAGGCTGTATCTACG 826
 DB 679 CACCTATGAGCAATGAAGTGTCTCAACCCGAGGATGTGGCGAGGCTGTATCTACG 620
 QY 827 TCCTCAGACCCCGCACACATCCAGATTGGAGACATCCAGATGAGGCCACCGAGCAGG 886
 DB 619 TCCTCAGACCCCGCACACATCCAGATTGGAGACATCCAGATGAGGCCACCGAGCAGG 560
 QY 887 TGACCTAGTACGTGGGAGCTCTCTCCCTCCCTCCCTCATGGTGGCTCTCTCTG 946
 DB 559 TGACCTAGTACGTGGGAGCTCTCTCTCCCTCCCTCATGGTGGCTCTCTCTG 500
 QY 947 CTCTGATTTTGGTGTGATTTCTGGATCAGGGATACCACTTCTCTGTCACACCCCGA 1006
 DB 499 CTCTGATTTTGGTGTGATTTCTGGATCAGGGATACCACTTCTCTGTCACACCCCGA 440
 QY 1007 CCAGGGGCTAGAAAATTTGTTGAGATTTTATATCATCTTGTCAAATTTGCTTCACTGT 1066
 |||||

Db 439 CCAGGGCTAGAAAATTTGTTTGAGATTTTATATCATCTTGTCAAAATGCTTCAGTTGT 380
Qy 1067 AAATGTGAAAATCGGCTGGGAAAGAGGTTGTCCTAAATGTTTACTTGTAACT 1126
Db 379 AAATGTGAAAATCGGCTGGGAAAGAGGTTGTCCTAAATGTTTACTTGTAACT 320
Qy 1127 TGTTCCTTGTGCCCTGGGCACTTGGCCCTTGTCTGCTCTCAGTGTCTTCCCTTTGACATG 1186
Db 319 TGTTCCTTGTGCCCTGGGCACTTGGCCCTTGTCTGCTCTCAGTGTCTTCCCTTTGACATG 260
Qy 1187 GGAAGGAGTTGTGGCCAAATCCCACTCTTCTTGCACCTCAAGCTGTGCTGCTCAGGGC 1246
Db 259 GGAAGGAGTTGTGGCCAAATCCCACTCTTCTTGCACCTCAAGCTGTGCTGCTCAGGGC 200
Qy 1247 TGGGCTGGCAGAGGAGGCTTCACTTATATCTGTGTGTATTCAGGGCTCCAGACTT 1306
Db 199 TGGGCTGGCAGAGGAGGCTTCACTTATATCTGTGTGTATTCAGGGCTCCAGACTT 140
Qy 1307 CCTCCTTGTGCTGCCCACTGACACCTCTCCCTCTTATCTATCTCTCTGCTGCCCTCCCA 1366
Db 139 CCTCCTTGTGCTGCCCACTGACACCTCTCCCTCTTATCTATCTCTCTGCTGCCCTCCCA 80
Qy 1367 GCGGAGTCTGCTTCTGCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1425
Db 79 GCGGAGTCTGCTTCTGCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 21

RESULT 4
LOCUS BQ278156 1071 bp mRNA linear EST 07-MAY-2002
DEFINITION AGENCOURT_7061694 NIH_MGC_109 Homo sapiens cDNA clone IMAGE:5805128
5', mRNA sequence.
ACCESSION BQ278156
VERSION BQ278156.1 GI:20488364
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLC2043 row: n column: 09
High quality sequence stop: 603.
Location/Qualifiers
1. .1071
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5805128"
/tissue_type="teratocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_109"
/notes="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

FEATURES

source
1. .1071
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5805128"
/tissue_type="teratocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_109"
/notes="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

Query Match 43.5%; Score 654; DB 13; Length 1071;
Best Local Similarity 100.0%; Pred. NO. 4.9e-212; Indels 0; Gaps 0;
Matches 654; Conservative 0; Mismatches 0;
Qy 17 GCAGGTGGCGGGCGGCGGAGAGGCGGCGGCGGTGAGTCCCTCGACCCCGTGTGG 76
Db 1 GCAGGTGGCGGGCGGCGGAGAGGCGGCGGCGGTGAGTCCCTCGACCCCGTGTGG 60
Qy 77 GCTAGTCAGGAGCGGCGGCGGCGGCGGCGGTGAGTCCCTCGACCCCGTGTGG 136
Db 61 GCTAGTCAGGAGCGGCGGCGGCGGCGGCGGTGAGTCCCTCGACCCCGTGTGG 120
Qy 137 GCGCGAGCGGCGGCGGCGGCGGCGGCGGTGAGTCCCTCGACCCCGTGTGG 196
Db 121 GCGCGAGCGGCGGCGGCGGCGGCGGCGGTGAGTCCCTCGACCCCGTGTGG 180
Qy 197 GCGCGCGGCGGCGGCGGCGGCGGCGGCGGTGAGTCCCTCGACCCCGTGTGG 256
Db 181 GCGCGCGGCGGCGGCGGCGGCGGCGGCGGTGAGTCCCTCGACCCCGTGTGG 240
Qy 257 AGGAGCTGGCTGCTGAATGTAAGAGTGAGGCTACCCCGGCGGTGAGTCCCTCG 316
Db 241 AGGAGCTGGCTGCTGAATGTAAGAGTGAGGCTACCCCGGCGGTGAGTCCCTCG 300
Qy 317 GTGACCTATCAATGAAGAGGACATCTCTCCATGTTCTCAGCTATCGTTCTCAG 376
Db 301 GTGACCTATCAATGAAGAGGACATCTCTCCATGTTCTCAGCTATCGTTCTCAG 360
Qy 377 GCGGTGTAGACATCTGCATCAACAATGTGGCGGCGGCGGTGAGTCCCTCG 436
Db 361 GCGGTGTAGACATCTGCATCAACAATGTGGCGGCGGCGGTGAGTCCCTCG 420
Qy 437 GCAGCACCGTGTGGAGGACATGTTCAATGTGAAGCTGCGGCGGCGGTGAGTCC 496
Db 421 GCAGCACCGTGTGGAGGACATGTTCAATGTGAAGCTGCGGCGGCGGTGAGTCC 480
Qy 497 CACGGAGAGCTTACCAGTCCATGAAGAGCGGAATGTGGAGCGATGGGCGACATCA 556
Db 481 CACGGAGAGCTTACCAGTCCATGAAGAGCGGAATGTGGAGCGATGGGCGACATCA 540
Qy 557 TCAATAGCATGTCTGGCCACCGAGTGTACCCCTGCTGTGTGACCCACCTTCTAT 616
Db 541 TCAATAGCATGTCTGGCCACCGAGTGTACCCCTGCTGTGTGACCCACCTTCTAT 600
Qy 617 CCAAGTATGCCGTCACTGGCTGACAGAGGAGCTGAGGCAAGAGCTTCGGGAGG 670
Db 601 CCAAGTATGCCGTCACTGGCTGACAGAGGAGCTGAGGCAAGAGCTTCGGGAGG 654

RESULT 5

LOCUS BQ278156 627 bp mRNA linear EST 06-MAR-2002
DEFINITION X-EST0127637 S13KWS5 Homo sapiens cDNA clone S13KWS5-43-A10 5',
mRNA sequence.
ACCESSION BQ278156
VERSION BQ278156.1 GI:19204411
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 627)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontiers Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470

ORIGIN

Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 43 row: A column: 10
High quality sequence stop: 627.

FEATURES

location/Qualifiers
1..627
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S13KMS-43-A10"
/tissue_type="myeloma"
/cell_line="KMS-5"
/lab_host="Top10F"
/clone_lib="S13KMS5"
/note="Vector: pCNS; Site 1: EcoRI; Site 2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tobacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoRI which site is treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Beig method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORIGIN

Query Match 41.7%; Score 627; DB 12; Length 627;
Best Local Similarity 100.0%; Pred. No. 9.2e-203;
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 AAGCAGGTGCGCGCGCGGAGAGAGCGCGCGCGTCTCAGCTCTCGACCCCGTGTG 74
DB 1 AAGCAGGTGCGCGCGCGGAGAGAGCGCGCGCGTCTCAGCTCTCGACCCCGTGTG 60

QY 75 GGGTAGTTCAGCGAGGCGGACGGCGCGCGTGGGCCATGCGCCAGCGCGCATGGAGCG 134
DB 61 GGGTAGTTCAGCGAGGCGGACGGCGCGCGTGGGCCATGCGCCAGCGCGCATGGAGCG 120

QY 135 GTGGCGGACCGGTGGCGGTGTCAGCGGGGCGCTCGGGGGGATCGCGGGCGCGGTGGC 194
DB 121 GTGGCGGACCGGTGGCGGTGTCAGCGGGGCGCTCGGGGGGATCGCGGGCGCGGTGGC 180

QY 195 CCGGGCCCTGTCCAGCAGGAGCTCAAGGTGGTGGGCTGCGCGCGCACTGTGGGCAACAT 254
DB 181 CCGGGCCCTGTTCAGCAGGAGCTCAAGGTGGTGGGCTGCGCGCGCACTGTGGGCAACAT 240

QY 255 CGAGGAGTGGCTCTGAATAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 314
DB 241 CGAGGAGTGGCTCTGAATAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 300

QY 315 ATGTGACCTATCAATGAAGAGGACATCTCTCCATGTTCTCAGCTATCCGTTCTCAGCA 374
DB 301 ATGTGACCTATCAATGAAGAGGACATCTCTCCATGTTCTCAGCTATCCGTTCTCAGCA 360

QY 375 CAGCGGTGTAGACATCTGCATCAACATGTGCGGTGGCGCGGCTGACACCCCTGCTCTC 434
DB 361 CAGCGGTGTAGACATCTGCATCAACATGTGCGGTGGCGCGGCTGACACCCCTGCTCTC 420

QY 435 AGGCAGCACAGTGGTGGAGGACATGTTCAATGTGAAGTGGCGCTCAGCATCTG 494
DB 421 AGGCAGCACAGTGGTGGAGGACATGTTCAATGTGAAGTGGCGCTCAGCATCTG 480

QY 495 CACACGGGAGCCCTACAGTCCATGAGGAGCGGAATGTGGACCATGGGACATCAATAA 554
DB 481 CACACGGGAGCCCTACAGTCCATGAGGAGCGGAATGTGGACCATGGGACATCAATAA 540

QY 555 CATCAATAGCATGTCTGGCCACCAGGTGTTACCCCTGTCTGTGACCCACTTCTATAGTGC 614

DB 541 CATCAATAGCATGTCTGGCCACCAGTGTACCCCTGTGTGACCCACTTCTATAGTGC 600
QY 615 CACCAAGTATGCCGTCACTGCGCTGAC 641
DB 601 CACCAAGTATGCCGTCACTGCGCTGAC 627

RESULT 6
BG741589 822 bp mRNA linear EST 15-MAY-2001
LOCUS 602635416F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4780659 5',
DEFINITION mRNA sequence.
ACCESSION BG741589
VERSION BG741589.1 GI:14052242
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 822)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10639 row: p column: 04
High quality sequence stop: 803.

FEATURES
source
1..822
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4780659"
/lab_host="DHL0B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Skn3"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Query Match 41.5%; Score 624; DB 12; Length 822;
Best Local Similarity 100.0%; Pred. No. 8.5e-202;
Matches 624; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 GTGGGCGCATGCGCGCGCGCGGATGAGCGGTGGCGGACCGGCTGGCGTGGTGAAGG 163
DB 94 GTGGGCGCATGCGCGCGCGCGGATGAGCGGTGGCGGACCGGCTGGCGTGGTGAAGG 153

QY 164 GGGCTCGGGGGGCGCATCGCGCGCGCGGCTGGCGGCGCTGGTCCAGCGGACTGAAGG 223
DB 154 GGGCTCGGGGGGCGCATCGCGCGCGCGGCTGGCGGCGCTGGTCCAGCGGACTGAAGG 213

QY 224 TGGTGGGCTCGGGCGCGCATCGTGGGCAACATCGAGGAGCTGGCTGCTGAATGAAGAGTG 283
DB 214 TGGTGGGCTCGGGCGCGCATCGTGGGCAACATCGAGGAGCTGGCTGCTGAATGAAGAGTG 273

QY 284 CAGGCTACCCCGGAGCTTTGATCCCTACAGATGTGACCTATCAATGAAGAGACATCC 343
DB 274 CAGGCTACCCCGGAGCTTTGATCCCTACAGATGTGACCTATCAATGAAGAGACATCC 333

QY 344 TCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTGTAGACATCTGCATCAACAATG 403
DB 334 TCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTGTAGACATCTGCATCAACAATG 393

404 CTGGCTTGGCCGGCCTGACACCTCTCTCAGCAGCAGCAGTGGTGGAGACATGT 463
 394 CTGGCTTGGCCGGCCTGACACCTCTCTCAGCAGCAGCAGTGGTGGAGACATGT 453
 464 TCAATGTGAACGTCTGGCCCTCAGCATCTGCACACGGGAAGCCTACAGTCCATGAAGG 523
 454 TCAATGTGAACGTCTGGCCCTCAGCATCTGCACACGGGAAGCCTACAGTCCATGAAGG 513
 524 AGCGGAATGTGGACGATGGGCACATCAATTAACATCAATPAGCATGTCTGGCCACCGAGTGT 583
 514 AGCGGAATGTGGACGATGGGCACATCAATTAACATCAATPAGCATGTCTGGCCACCGAGTGT 573
 584 TACCCCTCTCTGTGACCCACTTCTATAGTGCACCAAGTATGCCGTGCTGACAG 643
 574 TACCCCTCTCTGTGACCCACTTCTATAGTGCACCAAGTATGCCGTGCTGACAG 633
 644 AGGACTCAGCAAGAGCTTGGGAGGCCAGACCCACATCCGAGCCACGTGCACTCTCTC 703
 634 AGGACTCAGCAAGAGCTTGGGAGGCCAGACCCACATCCGAGCCACGTGCACTCTCTC 693
 704 CAGGTGTGGTGGAGACAAATTCG 727
 694 CAGGTGTGGTGGAGACAAATTCG 717

RESULT 7
 BM755987
 LOCUS
 DEFINITION K-EST0034121 S1SNUS Homo sapiens cDNA clone S1SNUS-38-A10 5', mRNA
 sequence.
 ACCESSION BM755987.1 GI:19085602
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 640)
 Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Kim,Y.S.
 21C Frontier Korean EST Project 2001
 Unpublished (2002)
 Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 38 row: A column: 10
 High quality sequence stop: 640.

FEATURES
 Location/Qualifiers
 1..640
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S1SNUS-38-A10"
 /sex="F"
 /tissue_type="Aescites"
 /cell_type="Lymphoblast-like"
 /cell_line="SNU-5"
 /lab_host="top10F"
 /clone_lib="S1SNUS"
 Note:Organ: Stomach; Vector: pCNS; Site:1: EcoRI;
 Site:2: NotI; The poly (A) + RNA was dephosphorylated with
 bacterial alkaline phosphatase (BAP) and then decapped
 with tobacco acid pyrophosphatase (TAP). The decapped
 intact mRNA was ligated with DNA-RNA linker including BcoR
 I site by treatment of T4 RNA ligase and the first strand
 cDNA was synthesized from oligo dt-selected mRNA by
 priming with dt-tailed vector. The dt-tailed vector was
 adjusted to have about 60nt. The cDNA vector was

ORIGIN
 Query Match 40.5%; Score 609; DB 12; Length 640;
 Best Local Similarity 100.0%; Pred. No. 1.2e-196;
 Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 350 TGTCTCAGCTATCCGTTCTCAGCAGCAGCGGTAGACATCTGCATCAACATGCTGGCT 409
 Db 1 TGTCTCAGCTATCCGTTCTCAGCAGCAGCGGTAGACATCTGCATCAACATGCTGGCT 60
 QY 410 TGGCCCGGCTGACACCCCTGCTCTCAGCAGCAGCAGTGGTGGAGACATGTTCAATG 469
 Db 61 TGGCCCGGCTGACACCCCTGCTCTCAGCAGCAGCAGTGGTGGAGACATGTTCAATG 120
 QY 470 TGAAGTGTCTGGCCCTCAGCATCTGCACACGGGAAGCCTACAGTCCATGAGGAGCGGA 529
 Db 121 TGAAGTGTCTGGCCCTCAGCATCTGCACACGGGAAGCCTACAGTCCATGAGGAGCGGA 180
 QY 530 ATGTGGACGATGGGCACATCAATTAACATCAATPAGCATGTCTGGCCACCGAGTGTACCCC 589
 Db 181 ATGTGGACGATGGGCACATCAATTAACATCAATPAGCATGTCTGGCCACCGAGTGTACCCC 240
 QY 590 TGTCTGTGACCCACTTCTATAGTGCACCAAGTATGCCGTCACTCGCTGACAGAGGAC 649
 Db 241 TGTCTGTGACCCACTTCTATAGTGCACCAAGTATGCCGTCACTCGCTGACAGAGGAC 300
 QY 650 TGAGCAAGAGCTTGGGAGGCCAGACCCACATCCGAGCCACGTGCACTCTCCAGTG 709
 Db 301 TGAGCAAGAGCTTGGGAGGCCAGACCCACATCCGAGCCACGTGCACTCTCCAGTG 360
 QY 710 TGGTGGAGACACAATTCGCTTCAAACTCCACGACCAAGACCCCTGAGAGGAGCGTGC 769
 Db 361 TGGTGGAGACACAATTCGCTTCAAACTCCACGACCAAGACCCCTGAGAGGAGCGTGC 420
 QY 770 CCTATGAGCAAAATGAAGTGTCTCAAAACCGAGGATGTGCCGAGGCTGTATCTAGCTCC 829
 Db 421 CCTATGAGCAAAATGAAGTGTCTCAAAACCGAGGATGTGCCGAGGCTGTATCTAGCTCC 480
 QY 830 TCAGACACCCCGCACACATCCAGATTGGAGACATCCAGATGAGGCCACCGAGGAGCGTGA 889
 Db 481 TCAGACACCCCGCACACATCCAGATTGGAGACATCCAGATGAGGCCACCGAGGAGCGTGA 540
 QY 890 CCTAGTGAATGTGGGAGCTCTCTTCCCTCCACCCCTTCATGGCTTGCCTTCCTCCTC 949
 Db 541 CCTAGTGAATGTGGGAGCTCTCTTCCCTCCACCCCTTCATGGCTTGCCTTCCTCCTC 600
 QY 950 TGGATTTTA 958
 Db 601 TGGATTTTA 609

RESULT 8
 BE617298
 LOCUS
 DEFINITION 601441804F1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3846168 5',
 mRNA sequence.
 ACCESSION BE617298
 VERSION BE617298.1 GI:9888236
 KEYWORDS
 SOURCE EST.
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 923)
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)

circularized with E. coli DNA ligase after digestion of
 EcoRI which site is also included in vector. An RNA strand
 converted to a DNA strand by Okayama-Berg method. The
 obtained cDNA vectors were used for transformation of
 competent cells E. coli Top10F' by electroporation method.
 The cDNA libraries constructed by this method are
 full-length enriched cDNA library."

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLN9558 row: 9 column: 01
High quality sequence stop: 698.
Location/Qualifiers
1. .923
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3846168"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 65"
/note="Organ: colon; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.8 Kb. Library constructed by Life
Technologies."

ORIGIN

Query Match 40.1%; Score 603; DB 10; Length 923;
Best Local Similarity 100.0%; Pred. No. 1.1e-194;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 332 AAGAGGACATCTCTCCATGTTCTCAGCTATCCGTTCTCAGCAGCGGTGAGACATCT 391
Db 1 AAGAGGACATCTCTCCATGTTCTCAGCTATCCGTTCTCAGCAGCGGTGAGACATCT 60

QY 392 GCATCAACAATGCTGGCTTGGCCGGCTGACACCTGCTCTCAGGAGCACCAGTGGTT 451
Db 61 GCATCAACAATGCTGGCTTGGCCGGCTGACACCTGCTCTCAGGAGCACCAGTGGTT 120

QY 452 GGAGGAGCATGTTCAATGTAACGTGCTGGCCCTCAGCATCTGCACGGAGCCTACC 511
Db 121 GGAGGAGCATGTTCAATGTAACGTGCTGGCCCTCAGCATCTGCACGGAGCCTACC 180

QY 512 AGTCCATGAGGAGCGGAATGTGACGATGGCAGCATCAATTAACATCAATGATGTCG 571
Db 181 AGTCCATGAGGAGCGGAATGTGACGATGGCAGCATCAATTAACATCAATGATGTCG 240

QY 572 GCCACCGAGTGTACCCCTGCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTCA 631
Db 241 GCCACCGAGTGTACCCCTGCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTCA 300

QY 632 CTGCGCTGACAGAGGACTGAGGCAAGAGCTTCGGAGGCCAGCCAGCATCCGAGCCA 691
Db 301 CTGCGCTGACAGAGGACTGAGGCAAGAGCTTCGGAGGCCAGCCAGCATCCGAGCCA 360

QY 692 CGTGCACTCTCCAGGTGTGGTGAGACACAATTCGCTTCAAACTCCACGACAGGACC 751
Db 361 CGTGCACTCTCCAGGTGTGGTGAGACACAATTCGCTTCAAACTCCACGACAGGACC 420

QY 752 CTGAGAGGCGAGCTGCCACCTATGAGCAATGAATGTCTCAAAACCGAGAGTGTGGCCG 811
Db 421 CTGAGAGGCGAGCTGCCACCTATGAGCAATGAATGTCTCAAAACCGAGAGTGTGGCCG 480

QY 812 AGGCTGTATCTAGTCTCTCAGCACCCTGAGCATCCAGATGGAGAGATCCAGATGA 871
Db 481 AGGCTGTATCTAGTCTCTCAGCACCCTGAGCATCCAGATGGAGAGATCCAGATGA 540

QY 872 GGCCCGAGGAGCGGTGACCTAGTGACTGTGGAGCTCCCTCTTCCCTCCCAACCTTCA 931
Db 541 GGCCCGAGGAGCGGTGACCTAGTGACTGTGGAGCTCCCTCTTCCCTCCCAACCTTCA 600

QY 932 TGG 934
|||

Db

601 TGG 603

RESULT 9

BM764342

LOCUS

DEFINITION

BM764342 602 bp mRNA linear EST 04-MAR-2002
X-EST0045842 S1SNUS Homo sapiens cDNA clone S1SNUS-39-A06 5', mRNA
sequence.

ACCESSION

BM764342

VERSION

BM764342.1

KEYWORDS

EST.

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 602)

AUTHORS

Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and Kim, Y.S.

TITLE

21C Frontier Korean EST Project 2001

JOURNAL

Unpublished (2002)

COMMENT

Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.krribb.re.kr

Plate: 39 row: A column: 06

High quality sequence stop: 602.

Location/Qualifiers

1. .602

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="S1SNUS-39-A06"

/sex="F"

/tissue_type="Ascites"

/cell_line="Lymphoblast-like"

/lab_host="Top10P"

/clone_lib="S1SNUS"

/note="Organ: Stomach; Vector: pCNS; Site 1: EcoRI;
Site 2: NotI; The poly (A) + RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tabacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10P by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."FEATURES
source

ORIGIN

Query Match 40.0%; Score 602; DB 12; Length 602;
Best Local Similarity 100.0%; Pred. No. 2.9e-194;
Matches 602; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 350 TGTTCCTCAGTATCCGTTCTCAGCAGCGGTGTAGACATCTGCATCAACAATGCTGGCT 409
Db 1 TGTTCCTCAGTATCCGTTCTCAGCAGCGGTGTAGACATCTGCATCAACAATGCTGGCT 60

QY 410 TGGCCCGGGCTGACACCCCTGCTCTCAGGAGCAGCAGTGTGTGGAGGAGCATGTTCAATG 469
Db 61 TGGCCCGGGCTGACACCCCTGCTCTCAGGAGCAGCAGTGTGTGGAGGAGCATGTTCAATG 120

QY 470 TGAACGTGTGGCCCTCAGCATCTGCACACGGGAAGCCTACCAGTCCATGAAGAGCGGA 529

Db

121

TGAACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCTTACAGTCCATGAAGAGCGGA

180

Qy

530

ATGTGGACGATGGGACACATCAATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCC

589

Db

181

ATGTGGACGATGGGACACATCAATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCC

240

Qy

590

TGCTGTGTGACCACTTCTATAGTCCACCAAGTATGCGCTCACTGCGCTGACAGGGAC

649

Db

241

TGCTGTGTGACCACTTCTATAGTCCACCAAGTATGCGCTCACTGCGCTGACAGGGAC

300

Qy

650

TGAGCAAGACGTTCGGAGGCGCCAGACCCACATCCGAGCCACGTGCACTCTCCAGTG

709

Db

301

TGAGCAAGACGTTCGGAGGCGCCAGACCCACATCCGAGCCACGTGCACTCTCCAGTG

360

Qy

710

TGGTGGAGACACAATTCGCGCTTCAAACTCCACGACCAAGGACCTGAGAAGGCGAGTGC

769

Db

361

TGGTGGAGACACAATTCGCGCTTCAAACTCCACGACCAAGGACCTGAGAAGGCGAGTGC

420

Qy

770

CCTATGAGCAAAATGAAGTGTCTCAAAACCGAGGATGTGGCGGAGCTGTATCTACGTCC

829

Db

421

CCTATGAGCAAAATGAAGTGTCTCAAAACCGAGGATGTGGCGGAGCTGTATCTACGTCC

480

Qy

830

TCAGCACCCCGCACACATCCAGATTGGAGACATCCAGATGAGGCCACGAGCAGGTGA

889

Db

481

TCAGCACCCCGCACACATCCAGATTGGAGACATCCAGATGAGGCCACGAGCAGGTGA

540

Qy

890

CCTAGTGACTGTGGAGCTCTCTCTTCCCTCCCGCCACCTTCATGCTTGCCTCTCGCTC

949

Db

541

CCTAGTGACTGTGGAGCTCTCTCTTCCCTCCCGCCACCTTCATGCTTGCCTCTCGCTC

600

Qy

950

TG 951

Db

601

TG 602

RESULT 10

AL559036

LOCUS

AL559036 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED

DEFINITION

Homo sapiens cDNA clone CS0DJ010YN13 5-PRIME, mRNA sequence.

ACCESSION

AL559036

VERSION

AL559036.2 GI:31283169

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

AUTHORS

Full-length cDNA libraries and normalization

TITLE

Unpublished (2001)

JOURNAL

On Feb 15, 2001 this sequence version replaced gi:12904138.

COMMENT

Contact: Genoscope

Genoscope, Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen. This sequence belongs to sequence cluster

8845.f For more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0DJ010CG07QPl&cluster=8845.f. Contact :

Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Paraday Avenue Genoscope sequence ID : CS0DJ010CG07QPl.

FEATURES

source

1..1022

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DJ010YN13"

/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"

/cell_line="JURKAT"

Qy

920

CCCCACCCCTTCATCGCTTGCCTCTGCTCTCGGATTTTAGTGTGTGATTTCTGATACAG

979

Db

367

CCCCACCCCTTCATCGCTTGCCTCTGCTCTCGGATTTTAGTGTGTGATTTCTGATACAG

426

Qy

980

GGATACACATTCCTGTGTCCACACCCCGACCGGCGCTAGAAAATTTGTTTGAGATTTTAT

1039

Db

427

GGATACACATTCCTGTGTCCACACCCCGACCGGCGCTAGAAAATTTGTTTGAGATTTTAT

486

Qy

1040

ATCATCTTGTCAAAATTTGCTTCAGTTGTAATGTGAAAAATGGGCTGGGGAAGGAGGTGG

1099

Db

487

ATCATCTTGTCAAAATTTGCTTCAGTTGTAATGTGAAAAATGGGCTGGGGAAGGAGGTGG

546

Qy

1100

TGTCCTCAATTTGTTTTTACTTGTAACTTGTGTTTGTGCCCCCTGGGCACTTGGCCTTTGTC

1159

Db

547

TGTCCTCAATTTGTTTTTACTTGTAACTTGTGTTTGTGCCCCCTGGGCACTTGGCCTTTGTC

606

Qy

1160

TGCTCTCAGTGTCTTCCCTTTGACATGCGGAAGAGTTGTGGCCAAAATCCCCATCTCT

1219

Db

607

TGCTCTCAGTGTCTTCCCTTTGACATGCGGAAGAGTTGTGGCCAAAATCCCCATCTCT

666

Qy

1220

TGCACCTCAACGTCTGTGCTCAGGCTGGGCTGCAGAGGAGGAGGCTTCACTTATATC

1279

Db

667

TGCACCTCAACGTCTGTGCTCAGGCTGGGCTGCAGAGGAGGAGGCTTCACTTATATC

726

Qy

1280

TGTGTGTTATCCAGGGCTCCAGACTTCTCTCTGCTGCCCCCACTGCACCTCTCCCC

1339

Db

727

TGTGTGTTATCCAGGGCTCCAGACTTCTCTCTGCTGCCCCCACTGCACCTCTCCCC

786

Qy

1340

CTTATCTATCTCTTCTGCTTCCCGACCGCAGTCTTGGCTTCTTGTGCCCCCTCCGGGT

1399

Db

787

CTTATCTATCTCTTCTGCTTCCCGACCGCAGTCTTGGCTTCTTGTGCCCCCTCCGGGT

846

Qy

1400

CATCCCTCCACTCTGACTCTGACTATGGCAGCAGAACACAGGCGCTGGCCAGTGGATT

1459

Db

847

CATCCCTCCACTCTGACTCTGACTATGGCAGCAGAACACAGGCGCTGGCCAGTGGATT

906

Qy

1460

TCATGGTGATCATTTAAAAAGAAAAATCGCAACCAAAAAA 1505

Db

907

TCATGGTGATCATTTAAAAAGAAAAATCGCAACCAAAAAA 952

RESULT 11

BM717452

LOCUS

BM717452 Homo sapiens

DEFINITION

UI-E-EJ0-ahn-c-08-0-UI.r1 UI-E-EJ0 Homo sapiens cDNA clone

ACCESSION

BM717452

VERSION

BM717452.1 GI:19030710

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

Bonaldi, M.F., Lennon, G. and Soares, M.B.

AUTHORS

Normalization and subtraction: two approaches to facilitate gene

TITLE

discovery

JOURNAL

Genome Res. 6 (9), 791-806 (1996)

MEDLINE

97044477

PUBMED

8889548

COMMENT

Contact: Soares, MB

/clone lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-Oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
The following repetitive elements were found in this cDNA
sequence: 603-644, >POLY A#Simple repeat

```

seq primer: M13 Reverse.
Location/Qualifiers
1. .649
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-EJ0-ahn-c-08-0-UI"
/tissue_type="fetal eyes, lens, eye anterior segment,
optic nerve, retina, Retina Foveal and Macular, RPE and
Choroid"
/dev_stage="fetal and adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-EJ0"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-EJ0 is a subtracted cDNA library constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tags for this library are: fetal eyes,
AGATCTCAGA; lens, CGATTAGCGA; eye anterior segment,
AATGCCGCAT; optic nerve, CCATTAGGTG; retina, CGCGG; Retina
Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This
library was created for the program, Gene Discovery in the
Visual System, supported by National Eye Institute (NEI)."
```

QY	872	GGCCACGGGACGAGTGACCTAGTAGACTGTGGGAGCTCCTCTTCCCTCCCGACCCCTTCA	931
Db	1	GGCCACGGGACGAGTGACCTAGTAGACTGTGGGAGCTCCTCTTCCCTCCCGACCCCTTCA	60
QY	932	TGGCTTGCCTCCCTCGCTCTGGATTTTAGTGTGTGATTTCTGGATCAGGGATACCACTTC	991
Db	61	TGGCTTGCCTCCCTCGCTCTGGATTTTAGTGTGTGATTTCTGGATCAGGGATACCACTTC	120
QY	992	CTGTCCACACCCCGACAGGGCTAGAAAATTTGTTTGAGATTTTATATCATCTTGTCAC	1051
Db	121	CTGTCCACACCCCGACAGGGCTAGAAAATTTGTTTGAGATTTTATATCATCTTGTCAC	180
QY	1052	AATTGCTTCAGTTGTAAATGTGAAAAATGGGCTGGGGAAGGAGTGCTGTCCTCAATTG	1111
Db	181	AATTGCTTCAGTTGTAAATGTGAAAAATGGGCTGGGGAAGGAGTGCTGTCCTCAATTG	240
QY	1112	TTTTACTTGTAACTTGTCTTGCCCTTGGGCACCTTGGCCCTTGTGCTCTCTCAGTGT	1171
Db	241	TTTTACTTGTAACTTGTCTTGCCCTTGGGCACCTTGGCCCTTGTGCTCTCTCAGTGT	300
QY	1172	CTTCCCTTTTGACATGGGAAGGAGTTGTGGCCAAAATCCCCATCTCTTGGACCTCAACG	1231
Db	301	CTTCCCTTTTGACATGGGAAGGAGTTGTGGCCAAAATCCCCATCTCTTGGACCTCAACG	360

[illegible]

RESULT 12	
BM674957/c	
LOCUS	
DEFINITION	650 bp mRNA linear EST 27-FEB-2002
ACCESSION	BM674957
VERSION	UI-E-EJ0-ahn-C-08-0-UI.s1
KEYWORDS	UI-E-EJ0 Homo sapiens cDNA clone
SOURCE	EST. mRNA sequence.
ORGANISM	BM674957.1 GI:18984855
REFERENCE	Homo sapiens (human)
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL	Bonaldi,M.F., Lennon,G. and Soares,M.B.
MEDLINE	Normalization and subtraction: two approaches to facilitate gene
PMEDID	discovery
COMMENT	Genome Res. 6 (9), 791-806 (1996)
	97044477
	8889548
	Contract: Soares MR

Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
The following repetitive elements were found in this cDNA
sequence: 1-38, >POLY A#Simple_repeat (matched complement)
Seq primer: M13 Forward
POLYA=Yes.

Location/Qualifiers

1. 650
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-EJO-ahn-c-08-0-UI"
/tissue_type="fetal eyes, lens, eye anterior segment,
optic nerve, retina, Retina Foveal and Macular, RPE and
Choroid"
/dev_stage="fetal and adult"
/lab_host="PH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-EJO"
/note="Organ: eye; Vector: p7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: Ecor I; Site 2: Not I;
UI-E-EJO is a subcloned cDNA library constructed

FEATURES source

according to Bonaldo, Lemmon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p773-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are: fetal eyes, AGAATCAAGA; lens, CGATTAGCA; eye anterior segment, AATGCCGAT; optic nerve, CCATTAAGTG; retina, CGCGG; Retina Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI). TAG TISSUE=human retina TAG LIB=UI-E-E70 TAG_SEQ=CCGGC"

ORIGIN		Query Match	38.6%; Score 581; DB 12; Length 650;
		Best Local Similarity	100.0%; Pred. No. 3.8e-187;
		Mismatches	581; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	925	CCCTTCATGGCTGCTCCTGCTCTGGATTTTAGTGTGATTTCTGGATCAGCGGATA	984
Db	587	CCCTTCATGGCTGCTCCTGCTCTGGATTTTAGTGTGATTTCTGGATCAGCGGATA	528
QY	985	CCACTTCCTGTCACACCCGACAGGGGCTAGAAAAATTTGTTGAGATTTTATATCAT	1044
Db	527	CCACTTCCTGTCACACCCGACAGGGGCTAGAAAAATTTGTTGAGATTTTATATCAT	468
QY	1045	CTTGTCAAATGCTTCAGTTGTAATGTGAAAAATGGGCTGGGAAAGAGGTGTGTCC	1104
Db	467	CTTGTCAAATGCTTCAGTTGTAATGTGAAAAATGGGCTGGGAAAGAGGTGTGTCC	408
QY	1105	CTAATGTTTACTTGTAACTTGTCTGCCCCCTGGGCACTTGGCCTTTGCTGCTC	1164
Db	407	CTAATGTTTACTTGTAACTTGTCTGCCCCCTGGGCACTTGGCCTTTGCTGCTC	348
QY	1165	TCAGTGTCTTCCCTTTGACATGGGAAAGAGTGTGGCAAAATCCCAATCTCTTGCAC	1224
Db	347	TCAGTGTCTTCCCTTTGACATGGGAAAGAGTGTGGCAAAATCCCAATCTCTTGCAC	288
QY	1225	CTCAACGCTGTGGCTCAGGGCTGGGGTGGCAGAGGAGGCTTACCTTATATCTGTGT	1284
Db	287	CTCAACGCTGTGGCTCAGGGCTGGGGTGGCAGAGGAGGCTTACCTTATATCTGTGT	228
QY	1285	TGTTATCCAGGGTCCAGACTTCTCCTCTGCTGCCCCCACTGCACCCCTCTCCCCCTAT	1344
Db	227	TGTTATCCAGGGTCCAGACTTCTCCTCTGCTGCCCCCACTGCACCCCTCTCCCCCTAT	168
QY	1345	CTATCTCTCTCTGGCTCCCGCAGCCAGCTTTGGCTTTCTGTCCCTCTCTGGGGTCATCC	1404
Db	167	CTATCTCTCTCTGGCTCCCGCAGCCAGCTTTGGCTTTCTGTCCCTCTCTGGGGTCATCC	108
QY	1405	CTCCACTCTGACTCTGACTATGTCACAGACACACAGGCGCTGGCCCTGGAATTCATG	1464
Db	107	CTCCACTCTGACTCTGACTATGTCACAGACACACAGGCGCTGGCCCTGGAATTCATG	48
QY	1465	GTGATCATTAATAAAGAAAAATCGCAACCAAAAAA 1505	
Db	47	GTGATCATTAATAAAGAAAAATCGCAACCAAAAAA 7	

RESULT 13
BU624717/c
LOCUS
DEFINITION
UI-H-FGI-bgi-h-13-0-UI.s1 NCI CGAP_FGI Homo sapiens cDNA clone
UI-H-FGI-bgi-h-13-0-UI 3', mRNA sequence.
BU624717
VERSION
BU624717.1 GI:23290932
KEYWORDS
EST.
Homo sapiens (human)
SOURCE

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 672)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: James Martin
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
Location/Qualifiers
1..672
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="UI-H-FGI-bgi-h-13-0-UI"
/tissue_type="Cell lines"
/dev_stage="Adult"
/lab_hosts="DH10B (Life Technologies)"
/clone_lib="NCI-CGAP_FGI"
/note="Organ: Enchondroma; Vector: p773-Pac (Pharmacia)
with a modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI CGAP FGI is a normalized cDNA library obtained from a
pool of mRNA from 2 cell lines from Enchondroma tissues.
The library was constructed according to Bonaldo, Lemmon
and Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated
to an EcoR I adaptor, digested with Not I, and cloned
directionally into p773-Pac vector. The oligonucleotide
used to prime the synthesis of first-strand cDNA contains
a library tag sequence that is located between the Not I
site and the (dT)18 tail. The sequence tag for this
library is CGGTCATC. The cell lines were provided by Dr.
James Martin from the University of Iowa.
TAG TISSUE=Enchondroma cell line (Mix of EN1 and EN2)
TAG LIB=UI-H-FGI
TAG_SEQ=CGGTCATC"

ORIGIN

Query Match	38.3%;	Score 576;	DB 13;	Length 672;
Best Local Similarity	100.0%;	Pred. No. 1.9e-185;		
Matches	576;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	915	TCCTCCCCACCTTCATGGCTTGCTCCTCGCTCTCGATTTTAGTGTGTGATTTCTCGGA	974	
Db	597	TCCTCCCCACCTTCATGGCTTGCTCCTCGCTCTCGATTTTAGTGTGTGATTTCTCGGA	538	
QY	975	TCACGGGATACCACTTCTCTGTCCACACCCCGACAGGGGCTAGAAAAATTTGTTGAGATT	1034	
Db	537	TCACGGGATACCACTTCTCTGTCCACACCCCGACAGGGGCTAGAAAAATTTGTTGAGATT	478	
QY	1035	TTTATATCATCTTGTCAAATGCTTCAGTTGTAATGTGAAAAATGGGCTGGGGAAGA	1094	
Db	477	TTTATATCATCTTGTCAAATGCTTCAGTTGTAATGTGAAAAATGGGCTGGGGAAGA	418	
QY	1095	GGTGGTGCCCTAATGTTTACTTGTAACTGTCTTGTGCGCCCTGGGCACATTGSCCT	1154	
Db	417	GGTGGTGCCCTAATGTTTACTTGTAACTGTCTTGTGCGCCCTGGGCACATTGSCCT	358	
QY	1155	TTGTCTGCTCTCAGTGTCTTCCCTTTGACATGGGAAAGAGTGTGTGGCCAAAATCCCCAT	1214	
Db	357	TTGTCTGCTCTCAGTGTCTTCCCTTTGACATGGGAAAGAGTGTGTGGCCAAAATCCCCAT	298	
QY	1215	CTTCTTTGCACCTCAACGCTGTGTGGCTGTGGGCTGGGGTGGCAGGGAGGGCTTCACCTT	1274	

```

Db      297  CTTCTTGCACCTCAACGCTGTGGCTCAGGGCTGGGGTGGCAGAGGAGGCGCTTCACTT 238
QY      1275  ATATCTGTCTTGTATCCAGGGCTCCAGACTTCTCTCTCTGCTGCTGCCACCTGACCCCTC 1334
Db      237  ATATCTGTGTGTATCCAGGGCTCCAGACTTCTCTCTCTGCTGCTGCCACCTGACCCCTC 178
QY      1335  TCCCTCTATCTATCTCTCTCTGCTGCCCTGCCAGCCAGCTTGTGGCTTCTGTGCTCCCTCT 1394
Db      177  TCCCTCTATCTATCTCTCTCTGCTGCCCTGCCAGCCAGCTTGTGGCTTCTGTGCTCCCTCT 118
QY      1395  GGGGTATCTCTCTCTCTCTGCTGCCCTGCCAGCCAGCTTGTGGCTTCTGTGCTCCCTCT 1454
Db      117  GGGGTATCTCTCTCTCTCTGCTGCCCTGCCAGCCAGCTTGTGGCTTCTGTGCTCCCTCT 58
QY      1455  GGATTTTCATGGTGATCATTAATAAAGAAAAATCGCA 1490
Db      57  GGATTTTCATGGTGATCATTAATAAAGAAAAATCGCA 22

RESULT 14
BU539642
LOCUS   AGENCOURT_10224058 NIH_MGC_107 Homo sapiens cDNA clone EST 13-SEP-2002
DEFINITION IMAGE:6570429 5', mRNA sequence.
ACCESSION BU539642
VERSION   BU539642.1 GI:22850083
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 937)
AUTHORS   NIH-MGC http://mgi.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-remail.nih.gov
          Tissue Procurement: ATCC
          cDNA Library Preparation: Rubin Laboratory
          CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Agencourt Bioscience Corporation
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLCM2759 row: e column: 21
          High quality sequence stop: 527.

FEATURES             Location/Qualifiers
     source            1..937
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="IMAGE:6570429"
                        /tissue_type="adenocarcinoma, cell line"
                        /lab_host="DH10B (phage-resistant)"
                        /clone_lib="NIH_MGC_107"
                        /note="Organ: breast; Vector: pOTB7; Site: 1: EcoRI;
                        Site 2: XhoI; cDNA made by oligo-dT priming.
                        Directionally cloned into EcoRI/XhoI sites using the
                        following 5' adaptor: GGCAGGAG(G). Library constructed by
                        Ling Hong in the laboratory of Gerald M. Rubin (University
                        of California, Berkeley) using ZAP-cDNA synthesis kit
                        (Stratagene) and Superscript II RT (Life Technologies).
                        Note: this is a NIH_MGC Library."

ORIGIN
Query Match      38.2%; Score 575; DB 13; Length 937;
Best Local Similarity 100.0%; Pred. No. 3.5e-185;
Matches 575; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      135  GTGCGCGACCGCTGGCGTGTGTGACGCGGGGCTCGGGGGGATCGCGCGCCGTGGC 194
Db      1    GTGCGCGACCGCTGGCGTGTGTGACGCGGGGCTCGGGGGGATCGCGCGCCGTGGC 60

```

```

QY      195  CCGGGCCCTGGTCCAGCAGGGACTGAAGTGTGGGCTCGGCCCGCACGTGTGGGCAACAT 254
Db      61    CCGGGCCCTGGTCCAGCAGGGACTGAAGTGTGGGCTCGGCCCGCACGTGTGGGCAACAT 120
QY      255  CGAGAGCTGGCTGTGAATGTAAGAGTGCAGGCTACCCCGGAGCTTTGATCCCTTACAG 314
Db      121  CGAGAGCTGGCTGTGAATGTAAGAGTGCAGGCTACCCCGGAGCTTTGATCCCTTACAG 180
QY      315  ATGTGACCTATCAAAATGAAGAGGACATCTCTCCATGTTCTCAGCTATCCGTTCTCAGCA 374
Db      181  ATGTGACCTATCAAAATGAAGAGGACATCTCTCCATGTTCTCAGCTATCCGTTCTCAGCA 240
QY      375  CAGCGGTGTAGACATCTGCATCAACAAATGCTGGCTTGGCCCGGCTGACACCTGTCTC 434
Db      241  CAGCGGTGTAGACATCTGCATCAACAAATGCTGGCTTGGCCCGGCTGACACCTGTCTC 300
QY      435  AGGCAGCACCAGTGGTGGAGGACATCTTCAATGTGAACGTGCTGGCCCTCAGCATCTG 494
Db      301  AGGCAGCACCAGTGGTGGAGGACATCTTCAATGTGAACGTGCTGGCCCTCAGCATCTG 360
QY      495  CACCGGAAGCCTACCACTGATCCATGAAGAGCGGAATGTGGACGATGGGCACATCATTA 554
Db      361  CACCGGAAGCCTACCACTGATCCATGAAGAGCGGAATGTGGACGATGGGCACATCATTA 420
QY      555  CATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGTGTGACCCACTTCTATAGTGC 614
Db      421  CATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGTGTGACCCACTTCTATAGTGC 480
QY      615  CACCAAGTATCCGCTCACTGCGTGACAGAGGAGTGAAGCAAGAGCTTCGGAGGCCCA 674
Db      481  CACCAAGTATCCGCTCACTGCGTGACAGAGGAGTGAAGCAAGAGCTTCGGAGGCCCA 540
QY      675  GACCCACATCCGAGCCACGTCATCTCTCCAGGTG 709
Db      541  GACCCACATCCGAGCCACGTCATCTCTCCAGGTG 575

```

```

RESULT 15
BM742038
LOCUS   K-EST0014773 S1SNU5s1 Homo sapiens cDNA clone S1SNU5s1-1-B09 5',
DEFINITION mRNA sequence.
ACCESSION BM742038
VERSION   BM742038.1 GI:19063367
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 573)
AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
          Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
          Kim,Y.S.
          21C Frontier Korean EST Project 2001
          Unpublished (2002)
          Contact: Kim YS
          Genome Research Center
          Korea Research Institute of Bioscience & Biotechnology
          52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
          Tel: +82-42-860-4470
          Fax: +82-42-860-4409
          Email: yongsung@mail.kribb.re.kr
          Plate: 1 row: B column: 09
          High quality sequence stop: 573.

FEATURES             Location/Qualifiers
     source            1..573
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="S1SNU5s1-1-B09"
                        /sex="F"
                        /tissue_type="Ascites"

```


/cell_type="Lymphoblast-like"
/cell_line="SNU-5"
/lab_host="Top10F"
/clone_lib="S1SNU5s1"
/note="Organ: Stomach; Vector: pCNS; Site 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library. After analyzing and
sequencing about 2,000 ~ 3,000 colonies in original cDNA
library, the abundant cDNAs were selected and amplified by
PCR reaction using vector region primer including T7
promotor as 5' primer and N(dT)14 as 3' primer. The PCR
products were used as template for synthesis of
biotinylated single stranded RNA by in vitro transcription
reaction. The synthesized cDNAs were hybridized with
antisense single stranded cDNAs prepared from original
library and incubated with avidin-gel. After removing
DNA-RNA hybrids by centrifuge, the substracted cDNA
libraries were constructed by transformation of the
remaining DNA into competent cells E. coli Top10F with
electroporation method."

ORIGIN

Query Match 38.1%; Score 573; DB 12; Length 573;
Best Local Similarity 100.0%; Pred. No. 2.1e-184; Mismatches 0; Gaps 0;
Matches 573; Conservative 0;

QY 350 TGTTCTCAGCTATCGGTTCTCAGCACAGCGGTGTAGACATCTGCATCAACAATGCTGGCT 409
DB 1 TGTTCTCAGCTATCGGTTCTCAGCACAGCGGTGTAGACATCTGCATCAACAATGCTGGCT 60

QY 410 TGGCCCGGCTGACACCCCTGCTCTCAGGCAGCACCAGTGGTGGGAAGGACATGTTCAATG 469
DB 61 TGGCCCGGCTGACACCCCTGCTCTCAGGCAGCACCAGTGGTGGGAAGGACATGTTCAATG 120

QY 470 TGAACGTGCTGGCCCTCAGCATCTGCACAGGAGCCTACCAGTCCATGAAGGCGGA 529
DB 121 TGAACGTGCTGGCCCTCAGCATCTGCACAGGAGCCTACCAGTCCATGAAGGCGGA 180

QY 530 ATGTGGACGATGGGCACATCAATTAATCAATAGCATGTCTGGCCACCGAGTGTATACCCC 589
DB 181 ATGTGGACGATGGGCACATCAATTAATCAATAGCATGTCTGGCCACCGAGTGTATACCCC 240

QY 590 TGTCTGTGACCACTTCTATAGTGCACCAAGATATCGGTCATCTGCCTGACAGGGAC 649
DB 241 TGTCTGTGACCACTTCTATAGTGCACCAAGATATCGGTCATCTGCCTGACAGGGAC 300

QY 650 TGAGGCAAGAGCTTCGGGAGGCCACACCATCCGAGCCAGTGCATCTCTCCAGGTG 709
DB 301 TGAGGCAAGAGCTTCGGGAGGCCACACCATCCGAGCCAGTGCATCTCTCCAGGTG 360

QY 710 TGTGTGAGACACAATTCGGCTTCAAACTCCACGAAGGACCTGTGAAGGCGAGCTGCCA 769
DB 361 TGTGTGAGACACAATTCGGCTTCAAACTCCACGAAGGACCTGTGAAGGCGAGCTGCCA 420

QY 770 CCTATGAGCAATGAAGTGTCTCAAAACCGAGATGTGCCGAGGCTGTATCTAGTCC 829
DB 421 CCTATGAGCAATGAAGTGTCTCAAAACCGAGATGTGCCGAGGCTGTATCTAGTCC 480

QY 830 TCAGCACCCCGCACACATCCAGATTGGAGACATCCAGATGAGGCCCCACCGAGCAGGTGA 889

Db 481 TCAGCACCCCGCACACATCCAGATTGGAGACATCCAGATGAGGCCCCACGAGCAGGTGA 540

QY 890 CCTAGTACTGTGGGAGCTCTCTCTTCCTCCCTCCC 922
DB 541 CCTAGTACTGTGGGAGCTCTCTCTTCCTCCCCTCCC 573

RESULT 16
BM744024 566 bp mRNA linear EST 01-MAR-2002
LOCUS K-EST0017362 S1SNU5s1 Homo sapiens cDNA clone S1SNU5s1-4-C01 5',
DEFINITION mRNA sequence.
ACCESSION BM744024
VERSION BM744024.1 GI:19065353
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 566)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kr.ibm.re.kr
Plate: 4 row: C column: 01
High quality sequence stop: 566.
Location/Qualifiers
1..566
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S1SNU5s1-4-C01"
/sex="F"
/tissue_type="Ascites"
/cell_type="Lymphoblast-like"
/lab_host="Top10F"
/clone_lib="S1SNU5s1"
/note="Organ: Stomach; Vector: pCNS; Site 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library. After analyzing and
sequencing about 2,000 ~ 3,000 colonies in original cDNA
library, the abundant cDNAs were selected and amplified by
PCR reaction using vector region primer including T7
promotor as 5' primer and N(dT)14 as 3' primer. The PCR
products were used as template for synthesis of
biotinylated single stranded RNA by in vitro transcription
reaction. The synthesized RNA probes were hybridized with
antisense single stranded cDNAs prepared from original
library and incubated with avidin-gel. After removing
DNA-RNA hybrids by centrifuge, the substracted cDNA
libraries were constructed by transformation of the

remaining DNA into competent cells E. coli Top10F' with electroporation method."

```
ORIGIN
Query Match      37.6%; Score 566; DB 12; Length 566;
Best Local Similarity 100.0%; Pred. No. 5e-182;
Matches 566; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 350 TGTCTCAGCATATCCGTTCTCAGCAGCGGTGTAGACATCTGCATCAACATGCTGGCT 409
Db 1 TGTCTCAGCATATCCGTTCTCAGCAGCGGTGTAGACATCTGCATCAACATGCTGGCT 60
QY 410 TGGCCGGCTGACACCCCTCTCTCAGGAGCAGCAGTGGTGGAGGACATGTTCAATG 469
Db 61 TGGCCGGCTGACACCCCTCTCTCAGGAGCAGCAGTGGTGGAGGACATGTTCAATG 120
QY 470 TGAAGCTGTGGCTCAGCATCTGCACACGGGAAGCCTACCATGTCATGAAGGAGCGGA 529
Db 121 TGAAGCTGTGGCTCAGCATCTGCACACGGGAAGCCTACCATGTCATGAAGGAGCGGA 180
QY 530 ATGTGGAGCATGGGCACATCAATCAATAGCATGTCTGCCACCGAGTGTACCCC 589
Db 181 ATGTGGAGCATGGGCACATCAATCAATAGCATGTCTGCCACCGAGTGTACCCC 240
QY 590 TGTCTGTGACCATCTTATAGTGCCACCAAGTATGCCGTCACTCGCTGACAGAGGAC 649
Db 241 TGTCTGTGACCATCTTATAGTGCCACCAAGTATGCCGTCACTCGCTGACAGAGGAC 300
QY 650 TGAGCAAGAGCTTCGGAGGCCAGACCCACATCGAGCCACATGTCCTCAGGTG 709
Db 301 TGAGCAAGAGCTTCGGAGGCCAGACCCACATCGAGCCACATGTCCTCAGGTG 360
QY 710 TGGTGGAGACAAATTCGCTTCAAACTCCAGCAAGGAGGAGTGTGGCGAGGCTGTATCTACGTCC 829
Db 361 TGGTGGAGACAAATTCGCTTCAAACTCCAGCAAGGAGGAGTGTGGCGAGGCTGTATCTACGTCC 420
QY 770 CCTATGAGCAATGAAGTGTCTCAAAACCGAGGATGTGGCGAGGCTGTATCTACGTCC 829
Db 421 CCTATGAGCAATGAAGTGTCTCAAAACCGAGGATGTGGCGAGGCTGTATCTACGTCC 480
QY 830 TCAGACCCCGCAGACATCAGATTTGAGACATCCAGATGAGGCCACCGAGCAGGTGA 889
Db 481 TCAGACCCCGCAGACATCAGATTTGAGACATCCAGATGAGGCCACCGAGCAGGTGA 540
QY 890 CCTAGTGAAGTGTGGAGTCTCTCTT 915
Db 541 CCTAGTGAAGTGTGGAGTCTCTCTT 566
```

```
RESULT 17
BG030249
LOCUS      602297547F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:4392011 5',
DEFINITION mRNA sequence.
ACCESSION BG030249
VERSION    BG030249.1 GI:12419347
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Email: cgabs-remail.nih.gov
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
```

http://image.llnl.gov
Plate: LLAMI0083 row: n column: 12
High quality sequence stop: 601.

FEATURES

source

```
/location="Qualifiers"
1. .932
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4392011"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_87"
/note="Organ: breast; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.383 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
```

ORIGIN

```
Query Match      36.1%; Score 543; DB 10; Length 932;
Best Local Similarity 100.0%; Pred. No. 2.6e-174;
Matches 543; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 TGCTGAATGTAAGAGTGAGGCTACCCCGGAGCTTTGATCCCTACAGATGTGACCTATC 326
Db 13 TGCTGAATGTAAGAGTGAGGCTACCCCGGAGCTTTGATCCCTACAGATGTGACCTATC 72
QY 327 AAATGAAGAGGACATCTCTCCATGTTCTCAGCTATCCGTTCTCAGCAGCGGTGTAGA 386
Db 73 AAATGAAGAGGACATCTCTCCATGTTCTCAGCTATCCGTTCTCAGCAGCGGTGTAGA 132
QY 387 CATCTGCATCAACAAATGCTGGCTTGGCCCGGCTGACACCTGCTCTCAGCAGCACCAG 446
Db 133 CATCTGCATCAACAAATGCTGGCTTGGCCCGGCTGACACCTGCTCTCAGCAGCACCAG 192
QY 447 TGGTTGGAAGGACATGTTCAATGTGAAGTGTGGCCCGGCTGACACCTGCTCTCAGCAGCACCAG 506
Db 193 TGGTTGGAAGGACATGTTCAATGTGAAGTGTGGCCCGGCTGACACCTGCTCTCAGCAGCACCAG 252
QY 507 CTACCACTCCATGAAGAGGAGCGAATGTGGAGCATGTGGGACATCAATCAATCAATAGCAT 566
Db 253 CTACCACTCCATGAAGAGGAGCGAATGTGGAGCATGTGGGACATCAATCAATCAATAGCAT 312
QY 567 GTCTGGCCACCGAGTGTACCCCTGTCTGTGACCCACTTCTATAGTCCACCAAGTATGC 626
Db 313 GTCTGGCCACCGAGTGTACCCCTGTCTGTGACCCACTTCTATAGTCCACCAAGTATGC 372
QY 627 CGTCACCTGCGCTGACAGAGGAGCTGAGGAGAGCTTGGGAGGCCCGACACCATCCG 686
Db 373 CGTCACCTGCGCTGACAGAGGAGCTGAGGAGAGCTTGGGAGGCCCGACACCATCCG 432
QY 687 AGCCACCTGATCTCTCCAGGTGTGGTGGAGACACAATTCGCCCTTCAAACTCCACGACAA 746
Db 433 AGCCACCTGATCTCTCCAGGTGTGGTGGAGACACAATTCGCCCTTCAAACTCCACGACAA 492
QY 747 GGACCCCTGAGAAGGAGCAGCTCCACCTATGAGCAAAATGAAGTGTCTCAAAACCCGAGGTGT 806
Db 493 GGACCCCTGAGAAGGAGCAGCTCCACCTATGAGCAAAATGAAGTGTCTCAAAACCCGAGGTGT 552
QY 807 GGC 809
Db 553 GGC 555
```

RESULT 18

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

```
BG231973
naf34g12.x1 Soares_NPBMC Homo sapiens cDNA clone IMAGE:4143166 3',
mRNA sequence.
BG231973
BG231973.1 GI:12727118
EST.
Homo sapiens (human)
```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 541)
AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: M. Bento Soares, Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D. and M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
info@image.llnl.gov
Seq primer: -400P from Gibco
High quality sequence stop: 480.
Location/Qualifiers
1. 541
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4143166"
/tissue_type="lymphocyte"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Soares NPBMC"
/note="Organ: blood; Vector: pT7D-Pac; Site 1: Not I;
Site 2: EcoRI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGGCCCGGGTTTTTTTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT73 vector. Library
is normalized; constructed in the laboratory of M. Bento
Soares (University of Iowa)."

FEATURES
source
1. 541
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S1SNUS-21-B02"
/sex="F"
/tissue_type="Ascites"
/cell_line="Lymphoblast-like"
/lab_host="Top10F"
/clone_lib="S1SNUS"
/note="Organ: Stomach; Vector: pcNS; Site 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then deacapped
with tabacco acid pyrophosphatase (TAP). The deacapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dT-selected mRNA by
priming with dT-tailed vector. The dT-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

ORIGIN
Query Match 35.9%; Score 541; DB 12; Length 541;
Best Local Similarity 100.0%; Pred. No. 1.6e-173; Indels 0; Gaps 0;
Matches 541; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 952 GATTTTAGTGTGTTCTGGATCAGCGGATACCACTTCCTGTCCACACCCCGACGAG 1011
DB 541 GATTTTAGTGTGTTCTGGATCAGCGGATACCACTTCCTGTCCACACCCCGACGAG 482
QY 1012 GCTAGAAAATTTGTTTGAGATTTTATATCATCTGTCAAAATTCAGTTGTAATG 1071
DB 481 GCTAGAAAATTTGTTTGAGATTTTATATCATCTGTCAAAATTCAGTTGTAATG 422
QY 1072 TCAAAAATGGCTGGGGAAGAGGTGGTGTCCCTAAATGTTTACTTGTAACTTGTTC 1131
DB 421 TGA AAAATGGCTGGGGAAGAGGTGGTGTCCCTAAATGTTTACTTGTAACTTGTTC 362
QY 1132 TTGTGCCCCCTGGGCACTTGGCCCTTGTCTGTCTCAGTGTCTCCCTTTGACATGGGAAA 1191
DB 361 TTGTGCCCCCTGGGCACTTGGCCCTTGTCTGTCTCAGTGTCTCCCTTTGACATGGGAAA 302
QY 1192 GGAGTTGTGGCCAAATCCCAATCTCTTTCGACCTCAAGCTGTGGCTCAGGCTGGGG 1251
DB 301 GGAGTTGTGGCCAAATCCCAATCTCTTTCGACCTCAAGCTGTGGCTCAGGCTGGGG 242
QY 1252 TGGCAGAGGAGGCGCTTACCTTATATCTGTGTGTATCCAGGGCTCCAGACTTCCCTCC 1311
DB 241 TGGCAGAGGAGGCGCTTACCTTATATCTGTGTGTATCCAGGGCTCCAGACTTCCCTCC 182
QY 1312 TCTGCTGCCCACTGCAGACCTCTCCCTTATCTATCTCTTCTGGCTCCCGACCCA 1371
DB 181 TCTGCTGCCCACTGCAGACCTCTCCCTTATCTATCTCTTCTGGCTCCCGACCCA 122
QY 1372 GTCTTGGCTTCTTGTCCCTCTGGGGTGCATCCCTCCACTCTGACTCTGACTATGGCAGC 1431

Db 121 GTCTTGGCTTCTTGTCCCTCTGGGGTGCATCCCTCCACTCTGACTCTGACTATGGCAGC 62
QY 1432 AGAACACAGGCGCTGGCCAGTGGATTTCATGTGATCATTAAGAAAGAAAAATCGAA 1491
DB 61 AGAACACAGGCGCTGGCCAGTGGATTTCATGTGATCATTAAGAAAGAAAAATCGAA 2
QY 1492 C 1492
DB 1 C 1
RESULT 19
BM737591
LOCUS K-EST0000094 S1SNUS Homo sapiens cDNA clone S1SNUS-21-B02 5', mRNA
DEFINITION sequence.
ACCESSION BM737591
VERSION BM737591.1 GI:19058920
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 582)
AUTHORS Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,
Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and
Kim, Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 21 row: B column: 02
High quality sequence stop: 582.
Location/Qualifiers
1. 582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S1SNUS-21-B02"
/sex="F"
/tissue_type="Ascites"
/cell_line="Lymphoblast-like"
/lab_host="Top10F"
/clone_lib="S1SNUS"
/note="Organ: Stomach; Vector: pcNS; Site 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then deacapped
with tabacco acid pyrophosphatase (TAP). The deacapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dT-selected mRNA by
priming with dT-tailed vector. The dT-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

ORIGIN
Query Match 35.3%; Score 531; DB 12; Length 582;
Best Local Similarity 99.8%; Pred. No. 3.8e-170; Indels 0; Gaps 0;
Matches 581; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongjung@mail.kribb.re.kr
Plate: 58 row: C column: 04
High quality sequence stop: 540.

FEATURES

Location/Qualifiers
1..540
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S13KMS5-58-C04"
/tissue_type="myeloma"
/cell_line="KMS-5"
/lab_host="Top10P"
/clone_lib="S13KMS5"
/note="Vector: PCNS; Site 1: EcoRI; Site 2: NotI; The poly (A) + RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tobacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoR I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10P' by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORIGIN

Query Match 34.3%; Score 516; DB 12; Length 540;
Best Local Similarity 100.0%; Pred. No. 4.9e-165;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 15 AAGCAGGTCTGGCGCGCGCAGGAGAGCGCGCGGCTCAGTCTCTCGACCCCGTGTCT 74
DB 1 AAGCAGGTCTGGCGCGCGCGCAGGAGAGCGCGCGGCTCAGTCTCTCGACCCCGTGTCT 60
QY 75 GGGCTAGTCTCAGCAGGCGGAGCGGGCGGGTGGGCCATGGCCAGGCGCGGATGGAGCG 134
DB 61 GGGCTAGTCTCAGCAGGCGGAGCGGGCGGGTGGGCCATGGCCAGGCGCGGATGGAGCG 120
QY 135 GTGGCGGACCGGCTGGGCGCTGTGACGGGGGCTCTGGGGGGCATCGGCGCGCGTGGC 194
DB 121 GTGGCGGACCGGCTGGGCGCTGTGACGGGGGCTCTGGGGGGCATCGGCGCGCGTGGC 180
QY 195 CGGGGCCCTGTGTCAGCAGGAGCTGAAGTGTGTGGGTGGCGCCCGCATGTGGGCAACAT 254
DB 181 CGGGGCCCTGTGTCAGCAGGAGCTGAAGTGTGTGGGTGGCGCCCGCATGTGGGCAACAT 240
QY 255 CGAGAGCTGGTGTCTGAATCTAAGTGTAGGCTACCCCGGACTTTGATCCCTACAG 314
DB 241 CGAGAGCTGGTGTCTGAATCTAAGTGTAGGCTACCCCGGACTTTGATCCCTACAG 300
QY 315 ATGTGACCTTCAATGAAGAGACATCTCTCCATGTTCTCAGCTATCCGTTCTCAGCA 374
DB 301 ATGTGACCTTCAATGAAGAGACATCTCTCCATGTTCTCAGCTATCCGTTCTCAGCA 360
QY 375 CAGCGGTGTAGACATCTGCATCAACATCTGGTGGCCCGGCTGACACCCCTGCTCTC 434
DB 361 CAGCGGTGTAGACATCTGCATCAACATCTGGTGGCCCGGCTGACACCCCTGCTCTC 420
QY 435 AGGCAGCACAGTGGTTGAAGACATGTTCAATGTGAACGTGTGGCCCTCAGCATCTG 494
DB 421 AGGCAGCACAGTGGTTGAAGACATGTTCAATGTGAACGTGTGGCCCTCAGCATCTG 480
QY 495 CACACGGGAGCGCTACCGTCCATGAAGAGCGGAA 530

Db 481 CACACGGGAGCGCTACCGTCCATGAAGAGCGGAA 516

RESULT 22

BM682228/c

LOCUS

DEFINITION

UI-E-EJ0-aio-j-09-0-UI.s1 UI-E-EJ0 Homo sapiens cDNA clone

UI-E-EJ0-aio-j-09-0-UI 3', mRNA sequence.

ACCSSION

BM682228

VERSION

BM682228.1 GI:18992124

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 572)

AUTHORS

Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE

Normalization and subtraction: two approaches to facilitate gene

discovery

JOURNAL

Genome Res. 6 (9), 791-806 (1996)

MEDLINE

97044477

PUBMED

COMMENT

Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman

cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com).

The following repetitive elements were found in this cDNA

sequence: 1-38, >POLY A#Simple_repeat (matched complement)

Seq primer: M13 Forward

POLYA=Yes.

Location/Qualifiers

1..572

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="UI-E-EJ0-aio-j-09-0-UI"

/tissue_type="fetal eyes, lens, eye anterior segment,

optic nerve, retina, Retina Foveal and Macular, RPE and

Choroid"

/dev_stage="fetal and adult"

/lab_host="DH10B (Life Technologies) (T1 phage resistant)"

/clone_lib="UI-E-EJ0"

/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a

modified polylinker; Site 1: EcoR I; Site 2: Not I;

UI-E-EJ0 is a subtracted cDNA library constructed

according to Bonaldo, Lennon and Soares, Genome Research,

6:791-806, 1996. First strand cDNA synthesis was primed

with an oligo-dT primer containing a Not I site. Double

stranded cDNA was ligated to an EcoR I adaptor, digested

with Not I, and cloned directionally into pT7T3-Pac

vector. The oligonucleotide used to prime the synthesis of

first-strand cDNA contains a library tag sequence that is

located between the Not I site and the (dT)18 tail. The

sequence tags for this library are: fetal eyes,

AGATCAGCA; lens, CGATTAGCGA; eye anterior segment,

AATGCCGAT; optic nerve, CCATTAAGTG; retina, CCGCG; Retina

Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This

library was created for the program, Gene Discovery in the

Visual System, supported by National Eye Institute (NEI).

TAG TISSUE=human retina

TAG_LIB=UI-E-EJ0

TAG_SEQ=CCGCG"

ORIGIN

```

Query Match      34.2%; Score 515; DB 12; Length 572;
Best Local Similarity 99.8%; Pred. No. 1e-164;
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 940 CTCCTGCTCTGGATTAGGTGTTGATTTCTGGATACGGATACCACTTCTCTGCCAC 999
DB |||||
572 CTCCTGCTCTGGATTAGGTGTTGATTTCTGGATACGGATACCACTTCTCTGCCAC 513
QY 1000 ACCCGACACAGGGCTAGAAAATTTGTTGAGATTTTATATCATCTTGTCAAAATGCTT 1059
DB |||||
512 ACCCGACACAGGGCTAGAAAATTTGTTGAGATTTTATATCATCTTGTCAAAATGCTT 453
QY 1060 CAGTTGTAATGTGAAAATGGCTGGGAAAGAGGTGGTCTCCCTAATGTTTACTT 1119
DB |||||
452 CAGTTGTAATGTGAAAATGGCTGGGAAAGAGGTGGTCTCCCTAATGTTTACTT 393
QY 1120 GTTAATCTTCTTCTGGCCCTGGGACACTTGGCCCTTGTCTGCTCTCAGTGTCTTCCCTT 1179
DB |||||
392 GTTAATCTTCTTCTGGCCCTGGGACACTTGGCCCTTGTCTGCTCTCAGTGTCTTCCCTT 333
QY 1180 TCACATGGGAAAGAGTGTGGCCAAATCCCATCTTCTTGCACCTCAAGCTCTGTGGC 1239
DB |||||
332 TCACATGGGAAAGAGTGTGGCCAAATCCCATCTTCTTGCACCTCAAGCTCTGTGGC 273
QY 1240 TCAGGCTGGGTCGACAGAGGAGGCTTTCACCTTATATCTGTGTGTTATCCAGGGCTC 1299
DB |||||
272 TCAGGCTGGGTCGACAGAGGAGGCTTTCACCTTATATCTGTGTGTTATCCAGGGCTC 213
QY 1300 CAGACTTCTCTCTGCTGCTGCCACCTGACACCTCTCCCTTATCTATCTCTCTCTCGG 1359
DB |||||
212 CAGACTTCTCTCTGCTGCTGCCACCTGACACCTCTCCCTTATCTATCTCTCTCTCGG 153
QY 1360 CTCCTCCAGCCAGCTTGGCTTCTTGTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1419
DB |||||
152 CTCCTCCAGCCAGCTTGGCTTCTTGTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 93
QY 1420 GACTATGGCAGCAGAACACAGGGGCTTGGCCAGTGGATTTTCATGGTGATCATTAATAAAA 1479
DB |||||
92 GACTATGGCAGCAGAACACAGGGGCTTGGCCAGTGGATTTTCATGGTGATCATTAATAAAA 33
QY 1480 GAAAAATCGCAACCAAAAAA 1505
DB |||||
32 GAAAAATCGCAACCAAAAAA 7

RESULT 23
BM743569
LOCUS
DEFINITION K-EST0016711 S1SNUS1 Homo sapiens cDNA clone S1SNUS1-5-C01 5',
mRNA sequence.
ACCESSION BM743569
VERSION BM743569.1 GI:19064898
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 550)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.krribb.re.kr
Plate: 5 row: C column: 01
High quality sequence stop: 550.
Location/Qualifiers

```

source

```

1..550
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S1SNUS1-5-C01"
/sex="P"
/tissue_type="Ascsites"
/cell_type="Lymphoblast-like"
/lab_host="Top10F"
/clone_lib="S1SNUS1"
/note="Organ: Stomach; Vector: pCNS; Site 1: EcoRI;
Site 2: NotI; The poly (A) RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tabacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library. After analyzing and
sequencing about 2,000 ~ 3,000 colonies in original cDNA
library, the abundant cDNAs were selected and amplified by
PCR reaction using vector region primer including T7
promotor as 5' primer and N(drf)14 as 3' primer. The PCR
products were used as template for synthesis of
biotinylated single stranded RNA by in vitro transcription
reaction. The synthesized RNA probes were hybridized with
antisense single stranded cDNAs prepared from original
library and incubated with avidin-gel. After removing
DNA-RNA hybrids by centrifuge, the subtracted cDNA
libraries were constructed by transformation of the
remaining DNA into competent cells E. coli Top10F, with
electroporation method."

```

ORIGIN

```

Query Match      34.0%; Score 511; DB 12; Length 550;
Best Local Similarity 100.0%; Pred. No. 2.4e-163;
Matches 511; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 390 CTGCATCAACAATGCTGGCTTGGCCGGCTGACACCTCTCTCAGGAGGACCACTGG 449
DB |||||
40 CTGCATCAACAATGCTGGCTTGGCCGGCTGACACCTCTCTCAGGAGGACCACTGG 99
QY 450 TTGGAAGGACATGTTCAATGTGAACGTGTGGCCCTCAGCATCTGCACACGGGAAGCTTA 509
DB |||||
100 TTGGAAGGACATGTTCAATGTGAACGTGTGGCCCTCAGCATCTGCACACGGGAAGCTTA 159
QY 510 CCAGTCCATGAAGAGCGGAATGTGACGATGGGCATCATTAACATCAATAGCATGTC 569
DB |||||
160 CCAGTCCATGAAGAGCGGAATGTGACGATGGGCATCATTAACATCAATAGCATGTC 219
QY 570 TGGCCACCGAGTGTACCCCTGTCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGT 629
DB |||||
220 TGGCCACCGAGTGTACCCCTGTCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGT 279
QY 630 CACTGCGCTGACAGAGGACTGAGGCAAGAGCTTCGGGAGGCGCCAGACCCACATCCGAGC 689
DB |||||
280 CACTGCGCTGACAGAGGACTGAGGCAAGAGCTTCGGGAGGCGCCAGACCCACATCCGAGC 339
QY 690 CAGTGCATCTCTCCAGGTGTGGTGAGACACAATTCGCTTCAAACTCCACGACAAGGA 749
DB |||||
340 CAGTGCATCTCTCCAGGTGTGGTGAGACACAATTCGCTTCAAACTCCACGACAAGGA 399
QY 750 CCTGTGAAGAGGAGCTGCCACCTATGAGCAATGAAGTGTCTCAAACTCCAGGATGTGC 809
DB |||||
400 CCTGTGAAGAGGAGCTGCCACCTATGAGCAATGAAGTGTCTCAAACTCCAGGATGTGC 459

```

QY	810	CGAGGCTGTTATCTACCTCTAGACACCCCGACACATCCAGATTGGAGATCCAGAT	869
Db	460	CGAGGCTGTTATCTACCTCTAGACACCCCGACACATCCAGATTGGAGATCCAGAT	519
QY	870	GAGGCCACGAGGAGGAGTGAAGTACTGACTG	900
Db	520	GAGGCCACGAGGAGGAGTGAAGTACTGACTG	550
RESULT 24			
LOCUS	BE386240	601273447F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3614770 5',	linear EST 21-JUL-2000
DEFINITION		mRNA sequence.	
ACCESSION	BE386240		
VERSION	BE386240.1	GI:9331605	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1	(bases 1 to 769)	
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC/DCFT/DTF cDNA Library Preparation: Ling Hong/Rubin Laboratory DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov Plate: L10M277 row: m column: 11 High quality sequence stop: 652.		
FEATURES	Location/Qualifiers		
source	1..769		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:3614770"		
	/tissue_type="melanotic melanoma"		
	/lab_host="DH10B (phage-resistant)"		
	/clone_lib="NIH_MGC_20"		
	/note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."		
ORIGIN			
Query Match	33.0%;	Score 496;	DB 10; Length 769;
Best Local Similarity	100.0%;	Pred. No. 2.6e-158;	
Matches 496;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	800	AGGATGTGGCCGAGGCTGTTATCTAGCTCCTCAGCACCCCGACACATCCAGATTGGAG	859
Db	1	AGGATGTGGCCGAGGCTGTTATCTAGCTCCTCAGCACCCCGACACATCCAGATTGGAG	60
QY	360	ACATCCAGATGAGGCCACGAGAGCGGTGACCTAGTGAAGTGGAGCTCTCTCCCT	919
Db	61	ACATCCAGATGAGGCCACGAGAGCGGTGACCTAGTGAAGTGGAGCTCTCTCCCT	120
QY	920	CCGACCCCTTCATGGCTTCGCTCCCTCGATTTGAGTTGATTTCTGGATCAGC	979
Db	121	CCGACCCCTTCATGGCTTCGCTCCCTCGATTTGAGTTGATTTCTGGATCAGC	180
QY	980	GGATACCACTTCTGTTCCACACCCCGACGAGGCGCTAGAAAAATTTGTTGAGATTTTAT	1039
Db	181	GGATACCACTTCTGTTCCACACCCCGACGAGGCGCTAGAAAAATTTGTTGAGATTTTAT	240
QY	1040	ATCATCTTGTCAAAATTCCTTCACTTGTAAATGTCAAAAATGGGTGGGAAAGGAGTGG	1059
Db	241	ATCATCTTGTCAAAATTCCTTCACTTGTAAATGTCAAAAATGGGTGGGAAAGGAGTGG	300
QY	1100	TGTCCTCAATTTGTTTACTTGTAACTTGTCTTGTGCCCTTGGGCACATTCGGCTTTGTC	1159
Db	301	TGTCCTCAATTTGTTTACTTGTAACTTGTCTTGTGCCCTTGGGCACATTCGGCTTTGTC	360
QY	1160	TGCTCTCAGTGTCTTCCCTTTGACATGGGAAAGAGTTGTGGCAAAATCCCATCTTCT	1219
Db	361	TGCTCTCAGTGTCTTCCCTTTGACATGGGAAAGAGTTGTGGCAAAATCCCATCTTCT	420
QY	1220	TGACCTCAACGCTGTGGCTCAGGCTGGGTGGCGAGAGGAGGCTTACCTTATATC	1279
Db	421	TGACCTCAACGCTGTGGCTCAGGCTGGGTGGCGAGAGGAGGCTTACCTTATATC	480
QY	1280	TGTGTTGTTATCCAGG	1295
Db	481	TGTGTTGTTATCCAGG	496
RESULT 25			
LOCUS	AT125225/c		
DEFINITION	q867f01.x1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1736473		
ACCESSION	AT125225		
VERSION	AT125225.1	GI:3593739	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1	(bases 1 to 489)	
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.		
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index		
JOURNAL	Unpublished (1997)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D. cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bbrp/image/image.html Insert Length: 598 Std Error: 0.00 Seq primer: -40ml3 fwd. ET from Amersham High quality sequence stop: 465.		
FEATURES	Location/Qualifiers		
source	1..489		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:1736473"		
	/sex="male"		
	/lab_host="DH10B"		
	/clone_lib="Soares_testis_NHT"		
	/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5]. TGTATCAATTCGAGTGGAGCGCGCCGCCCAATTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo."		
ORIGIN			


```
Query Match      32.5%; Score 489; DB 9; Length 489;
Best Local Similarity 100.0%; Pred. No. 7.6e-156;
Matches 489; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1013 GCTAGAAAATTTGTTGAGATTTTATATCATCTTGTCAAAATGCTTCAGTTGTAATGT 1072
Db 489 GCTAGAAAATTTGTTGAGATTTTATATCATCTTGTCAAAATGCTTCAGTTGTAATGT 430

QY 1073 GAAATATGGCTGGCGAAGAGGCTGCTCCCTAAATGTTTACTTGTAACTTGTCT 1132
Db 429 GAAATATGGCTGGCGAAGAGGCTGCTCCCTAAATGTTTACTTGTAACTTGTCT 370

QY 1133 TGTGCCCCCTGGGCACCTTGGCCCTTCTGCTCTCAGTGTCTTCCCTTTGACATGGGAAG 1192
Db 369 TGTGCCCCCTGGGCACCTTGGCCCTTCTGCTCTCAGTGTCTTCCCTTTGACATGGGAAG 310

QY 1193 GAGTTGTGGCCAAATATCCCATCTTCTTTCACCTCAACGCTCTGTGGCTCAGGGCTGGGGT 1252
Db 309 GAGTTGTGGCCAAATATCCCATCTTCTTTCACCTCAACGCTCTGTGGCTCAGGGCTGGGGT 250

QY 1253 GGCAGAGGAGGCTTACCTTATATCTGTGTTATTCAGGGCTCCAGACTTCTCTCT 1312
Db 249 GGCAGAGGAGGCTTACCTTATATCTGTGTTATTCAGGGCTCCAGACTTCTCTCT 190

QY 1313 CTGCTGCCCCCACTGCACCCCTCTCCCCCTTATCTATCTCTTCTGGCTCCCGCCAG 1372
Db 189 CTGCTGCCCCCACTGCACCCCTCTCCCCCTTATCTATCTCTGCTCCCGCCAG 130

QY 1373 TCTTGGCTTCTTGTCCCTCTCGGGGTCACTCCCTCACTCTGACTCTGACTATGCGACGA 1432
Db 129 TCTTGGCTTCTTGTCCCTCTCGGGGTCACTCCCTCACTCTGACTCTGACTATGCGACGA 70

QY 1433 GAACACAGGCGCTGGCCAGTGGATTCATGGTGATCATTAATAAGAAAAATCGCAAC 1492
Db 69 GAACACAGGCGCTGGCCAGTGGATTCATGGTGATCATTAATAAGAAAAATCGCAAC 10

QY 1493 CAAAAAAA 1501
Db 9 CAAAAAAA 1

RESULT 26
CA424831/c
LOCUS
DEFINITION
UI-H-FEI-bdw-e-17-0-UI.s1 NCI CGAP FEI Homo sapiens cDNA clone
UI-H-FEI-bdw-e-17-0-UI 3', mRNA sequence.
ACCESSION
CA424831
VERSION
EST.
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 566)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: James Martin
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
Location/Qualifiers
source
1..566
/organism="Homo sapiens"
/mol_type="mRNA"
```

```
/db_xref="taxon:9606"
/clone="UI-H-FEI-bdw-e-17-0-UI"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP FEI"
/note="Organ: Chondrosarcoma; Vector: pTT3-Pac
(Pharmacia) with a modified polylinker; Site 1: Ecor I;
Site 2: Not I; NCI CGAP FEI is a normalized cDNA library
derived from a pool of mRNA obtained from 3 cell lines
from grade II chondrosarcoma tissues. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an Ecor I
adaptor, digested with Not I, and cloned directionally
into pTT3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
CGTACCGAC. The cell lines were provided by Dr James
Martin from the University of Iowa.
TAG_TISSUE=Human grade 2 chondrosarcoma cell line pool
TAG_LIB=UI-H-FEI
TAG_SEQ=CGTACCGAC"
```

ORIGIN

```
Query Match      32.5%; Score 489; DB 14; Length 566;
Best Local Similarity 100.0%; Pred. No. 7.1e-156;
Matches 489; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 952 GATTTTGGTGTGATTTCTGGATCAGGGATACCACTTCCTGTCACACCCGACCAAG 1011
Db 566 GATTTTGGTGTGATTTCTGGATCAGGGATACCACTTCCTGTCACACCCGACCAAG 507

QY 1012 GGCTAGAAAATTTGTTGAGATTTTATATCATCTTGTCAAATGCTTCAGTTGTAATG 1071
Db 506 GGCTAGAAAATTTGTTGAGATTTTATATCATCTTGTCAAATGCTTCAGTTGTAATG 447

QY 1072 TGAATAATGGCTGGGGAAGAGGCTGTGTCCTTAATGTTTACTTGTAACTTGTTC 1131
Db 446 TGAATAATGGCTGGGGAAGAGGCTGTGTCCTTAATGTTTACTTGTAACTTGTTC 387

QY 1132 TTGTGCCCCCTGGGCACCTTGGCCCTTGTCTCTCTCAGTGTCTTCCCTTTGACATGGGAAA 1191
Db 386 TTGTGCCCCCTGGGCACCTTGGCCCTTGTCTCTCTCAGTGTCTTCCCTTTGACATGGGAAA 327

QY 1192 GGAGTTGTGGCCAAATATCCCATCTTCTTGACACTCAACGCTCTGTGGCTCAGGGCTGGG 1251
Db 326 GGAGTTGTGGCCAAATATCCCATCTTCTTGACACTCAACGCTCTGTGGCTCAGGGCTGGG 267

QY 1252 TGGCAGAGGAGGCTTCACTTATATCTGTGTTGTTATCCAGGGCTCCAGACTTCTCTCC 1311
Db 266 TGGCAGAGGAGGCTTCACTTATATCTGTGTTGTTATCCAGGGCTCCAGACTTCTCTCC 207

QY 1312 TCTGCTGCCCCACTGCACCCCTCTCCCTCTTATCTATCTCTCTCGGTCCCGACGCCA 1371
Db 206 TCTGCTGCCCCACTGCACCCCTCTCCCTCTTATCTATCTCTCTCGGTCCCGACGCCA 147

QY 1372 GTCTTGGCTTCTTGTCCCTCTCGGGGTATCCCTCACTCTGACTCTGACTATGCGAGC 1431
Db 146 GTCTTGGCTTCTTGTCCCTCTCGGGGTATCCCTCACTCTGACTCTGACTATGCGAGC 87

QY 1432 AGAACACCA 1440
Db 86 AGAACACCA 78
```

```
RESULT 27
BM723373
LOCUS
DEFINITION
UI-E-EJO-aio-j-09-0-UI.r1 UI-E-EJO Homo sapiens cDNA clone
UI-E-EJO-aio-j-09-0-UI 5', mRNA sequence.
```

```
BM723373      658 bp      mRNA      linear      EST 01-MAR-2002
UI-E-EJO-aio-j-09-0-UI.r1 UI-E-EJO Homo sapiens cDNA clone
UI-E-EJO-aio-j-09-0-UI 5', mRNA sequence.
```


ORIGIN

375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA Tel: 319 335 8250 Fax: 319 335 9565 Email: bentso-soares@uiowa.edu Tissue Procurement: Dr. Gregg Hageman cDNA Library preparation: Dr. M. Bento Soares, University of Iowa cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com). The following repetitive elements were found in this cDNA sequence: 1-38, >POLY_A#Simple_repeat (matched complement) Seq primer: M13 Forward POLYA=Yes.		1. 570 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="UI-E-EJ0-ahf-o-21-0-UI" /tissue="UI-E-EJ0-ahf-o-21-0-UI" optic nerve, retina, Retina Foveal and Macular, RPE and Choroid" /dev_stage="fetal and adult" /lab_host="DH10B (Life Technologies) (T1 phage resistant)" /clone_lib="UI-E-EJ0" /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-EJ0 is a subtracted cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are: fetal eyes, AGATCAAGA; lens, CGATTAGCA; eye anterior segment, AATGCCGAT; optic nerve, CCATTAGTG; retina, CCGCG; Retina Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NET). TAG_TISSUE=human retina TAG_LIB=UI-E-EJ0 TAG_SEQ=CCGCG"	
FEATURES		31.2%; Score 469; DB 12; Length 570; Query Match 100.0%; Pred. No. 4.4e-149; Indels 0; Gaps 0; Matches 469; Conservative 0; Mismatches 0;	
source		ORIGIN	
951 GGATTTTAGTGTGTTGATTTCTGATCAGCGGATACACCTTCCTGTCCACACCCCGACCCAG 1010		QY	
560 GGATTTTAGTGTGTTGATTTCTGATCAGCGGATACACCTTCCTGTCCACACCCCGACCCAG 501		Db	
1011 GGCTAGAAAAATTGTTGAGATTTTATATACATCTTGTCATATGCTTCAGTTGTAAT 1070		QY	
500 GGCTAGAAAAATTGTTGAGATTTTATATATATATATATATATATATATATATATATATAT 441		Db	
1071 GTGAAAATGGCGTGGGAAAGAGGTGTGTCCTAAATGTTTACTGTTAACTTGT 1130		QY	
440 GTGAAAATGGCGTGGGAAAGAGGTGTGTCCTAAATGTTTACTGTTAACTTGT 391		Db	
1131 CTTGTGCCCCCTGGGCACCTTGGCCTTTGTCTGCTCTCAGTGTCTTCCCTTTGACATGGGAA 1190		QY	
380 CTTGTGCCCCCTGGGCACCTTGGCCTTTGTCTGCTCTCAGTGTCTTCCCTTTGACATGGGAA 321		Db	
1191 AGGAGTTGTGGCAAAATCCCAATCTTCTTGCACTCAAGCTGTGGCTCAGGGCTGGG 1250		QY	
320 AGGAGTTGTGGCAAAATCCCAATCTTCTTGCACCTCAACGTGTGGCTCAGGGCTGGG 261		Db	
1251 GTGCAGAGGGAGGCGCTTCACCTTATATCTGTGTGTTATCCAGGGCTCCAGACTTCCTC 1310		QY	
RESULT 31		ORIGIN	
BM793929		Query Match 30.4%; Score 457; DB 12; Length 691;	
LOCUS K-EST0074920 S22SNU16n1 Homo sapiens cDNA clone S22SNU16n1-32-C08		Best Local Similarity 100.0%; Pred. No. 4.8e-145;	
DEFINITION 5', mRNA sequence.		Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
ACCESSION BM793929		QY	
VERSION BM793929.1 GI:19142161		Db	
KEYWORDS EST.		QY	
SOURCE Homo sapiens (human)		Db	
ORGANISM Homo sapiens		QY	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		Db	
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		QY	
REFERENCE 1 (bases 1 to 691)		Db	
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.		QY	
TITLE 21C Frontier Korean EST Project 2001		Db	
JOURNAL Unpublished (2002)		QY	
COMMENT Contact: Kim YS		Db	
Genome Research Center		QY	
Korea Research Institute of Bioscience & Biotechnology		Db	
52 Eoeun-gong Yuseong-gu, Daejeon 305-333, South Korea		QY	
Tel: +82-42-860-4470		Db	
Fax: +82-42-860-4409		QY	
Email: yongsung@mail.kribb.re.kr		Db	
Plate: 32 row: C column: 08		QY	
High quality sequence stop: 691.		Db	
Location/Qualifiers		QY	
1. 691		Db	
/organism="Homo sapiens"		QY	
/mol_type="mRNA"		Db	
/db_xref="taxon:9606"		QY	
/clone="S22SNU16n1-32-C08"		Db	
/sex="F"		QY	
/tissue_type="Ascites"		Db	
/cell_type="Lymphoblast-like"		QY	
/cell_line="SNU-16"		Db	
/lab_host="DH10B"		QY	
/clone_lib="S22SNU16n1"		Db	
/note="Organ: Stomach; Vector: pT7T3-Pac; Site 1: EcoRI; Site 2: NotI; The S22SNU16 library was contributed by the Soares laboratory and it was constructed as described by Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome Research 6(9): 791-806. RNA was prepared from harvested cells of SNU-16 culture. SNU-16 cell was obtained from Korean Cell Line Bank (KCLB). SNU-16 was established from ascitic fluids of Korean patients by Park J.G. et al. (1990), Cancer Res 50: 2773-2780."		QY	
786 GTGTCTCTCAACCCGAGGATGTGGCCGAGGCTGTATCTACGTCCTCAGCACCCCGCACA 845		Db	
222 GTGTCTCTCAACCCGAGGATGTGGCCGAGGCTGTATCTACGTCCTCAGCACCCCGCACA 281		QY	
846 CATCCAGATTGGAGACATCCAGATGAGGCCCGAGGAGGTGACCTAGTACTGTGGGA 905		Db	
282 CATCCAGATTGGAGACATCCAGATGAGGCCCGAGGAGGTGACCTAGTACTGTGGGA 341		QY	

Query Match 30.2%; Score 455; DB 10; Length 455;
Best Local Similarity 100.0%; Pred. No. 2.7e-144;

Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 The following repetitive elements were found in this cDNA
 sequence: 127-239, >GC-rich#Low_complexity
 Seq primer: M13 Reverse.
 Location/Qualifiers
 1. .516
 /organism="Homo sapiens"

mol_type="mRNA"
db_xref="taxon:9606"
/clone="UI-E-C11-afv-m-21-0-UI"
/tissue_type="RPE and Choroid"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-C11"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-C11 is a normalized cDNA library containing the following tissue(s): RPE and Choroid. The library was constructed according to Bonaldo, Lemon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

REFERENCE 1 (bases 1 to 886)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rgs@bbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM9753 row: n column: 02
High quality sequence stop: 589.
Location/Qualifiers
1. 886
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3921217"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 72"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."

FEATURES
source

ORIGIN
Query Match 30.1%; Score 453; DB 12; Length 516;
Best Local Similarity 100.0%; Pred. No. 1.2e-143;
Matches 453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCGGATCGACCCAGAGGTTCGGCGGGGGGCGGAGGAGCGCGGGGTTCAGCTCC 60
DB 38 CCGGATCGACCCAGAGGTTCGGCGGGGGGCGGAGGAGCGCGGGGTTCAGCTCC 97
QY 61 TCGACCCCGGTTCGGGTAGTCCAGCGAGGCGGAGCGGCGGCGGTTCAGCTCC 120
DB 98 TCGACCCCGGTTCGGGTAGTCCAGCGAGGCGGAGCGGCGGCGGTTCAGCTCC 157
QY 121 CCGGATCGAGCGGTTCGGGTAGTCCAGCGAGGCGGAGCGGCGGCGGTTCAGCTCC 180
DB 158 CCGGATCGAGCGGTTCGGGTAGTCCAGCGAGGCGGAGCGGCGGCGGTTCAGCTCC 217
QY 181 GCGCGGCGGTTCGGGCGGTTCGGGTAGTCCAGCGAGGCGGAGCGGCGGCGGTTCAGCTCC 240
DB 218 GCGCGGCGGTTCGGGCGGTTCGGGTAGTCCAGCGAGGCGGAGCGGCGGCGGTTCAGCTCC 277
QY 241 ACTGTGGGCAACATCGAGAGGTTCGGGTAGTCCAGCGAGGCGGAGCGGCGGCGGTTCAGCTCC 300
DB 278 ACTGTGGGCAACATCGAGAGGTTCGGGTAGTCCAGCGAGGCGGAGCGGCGGCGGTTCAGCTCC 337
QY 301 TTGATCCCTTACAGATGTGACCTATCAAAATGAAGAGGACATCTCTCCATGTTCTCAGCT 360
DB 338 TTGATCCCTTACAGATGTGACCTATCAAAATGAAGAGGACATCTCTCCATGTTCTCAGCT 397
QY 361 ATCCGTTCTCAGCAGCGGTTCAGATCTGCATCAAAATGAAGAGGACATCTCTCCATGTTCTCAGCT 420
DB 398 ATCCGTTCTCAGCAGCGGTTCAGATCTGCATCAAAATGAAGAGGACATCTCTCCATGTTCTCAGCT 457
QY 421 GACACCCCTGCTTCAGGAGCGACAGTGGTTGG 453
DB 458 GACACCCCTGCTTCAGGAGCGACAGTGGTTGG 490

RESULT 34
BE895089 886 bp mRNA linear EST 20-OCT-2000
LOCUS G01436007F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3921217 5',
DEFINITION mRNA sequence.
ACCESSION BE895089
VERSION BE895089.1 GI:10358131
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 886)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: rgs@bbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM9753 row: n column: 02
High quality sequence stop: 589.
Location/Qualifiers
1. 886
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3921217"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 72"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."

ORIGIN
Query Match 30.0%; Score 451; DB 10; Length 886;
Best Local Similarity 100.0%; Pred. No. 4.6e-143;
Matches 451; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 476 TGCTGGCCCTCAGCATCTGCACACGGGAAGCTACAGTCCATGAAGAGCGGAATGTGG 535
DB 1 TGCTGGCCCTCAGCATCTGCACACGGGAAGCTACAGTCCATGAAGAGCGGAATGTGG 60
QY 536 ACGATGGGCGACATCAATTAACATCAATAGATGTCTGGCCACCGAGTGTACCCCTGTCTG 595
DB 61 ACGATGGGCGACATCAATTAACATCAATAGATGTCTGGCCACCGAGTGTACCCCTGTCTG 120
QY 596 TGACCCACTTCTATAGTCCACCAAGTATCGCTCACTCGCTGACAGAGGACTGAGCG 655
DB 121 TGACCCACTTCTATAGTCCACCAAGTATCGCTCACTCGCTGACAGAGGACTGAGCG 180
QY 656 AAGAGCTTGGGGAGGCCAGACCCACATCCGAGCCACGTCGATCTCTCCAGGTGTGGTGG 715
DB 181 AAGAGCTTGGGGAGGCCAGACCCACATCCGAGCCACGTCGATCTCTCCAGGTGTGGTGG 240
QY 716 AGACACATTCGCTTCAAACTCCACGAGGAGCCCTGAGAGGCGAGCTGCCACCTATG 775
DB 241 AGACACATTCGCTTCAAACTCCACGAGGAGCCCTGAGAGGCGAGCTGCCACCTATG 300
QY 776 ASCAATGAAGTGTCTCAAAACCCGAGGATGTGGCCGAGGCTGTATCTACGCTCTCAGCA 835
DB 301 AGCAATGAAGTGTCTCAAAACCCGAGGATGTGGCCGAGGCTGTATCTACGCTCTCAGCA 360
QY 836 CCCCCGCACACATCCAGATTGGAGACATCCAGATGAGGCCCGAGGAGGAGGAGGAGGAGT 895
DB 361 CCCCCGCACACATCCAGATTGGAGACATCCAGATGAGGCCCGAGGAGGAGGAGGAGGAGT 920
QY 896 GACTGTGGGAGTCT 926
DB 421 GACTGTGGGAGTCT 451

RESULT 35
BE8973434 481 bp mRNA linear EST 18-DEC-1999
LOCUS w19a10.x1 Soares_Dieckgraefe_colon_NHUC Homo sapiens cDNA clone
DEFINITION IMAGE:2351034 3', mRNA sequence.

```

ACCESSION   AI673434
VERSION     AI673434.1  GI:4853165
KEYWORDS    EST.
SOURCE      Homo sapiens
ORGANISM    Homo sapiens (human)
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 481)
            NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE       Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            This clone is available royalty-free through LNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 579 Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 457.
FEATURES    source
            1..481
                Location/Qualifiers
                1..481
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:2351034"
                    /tissue_type="colonic mucosa from 5 ulcerative colitis
                    patients"
                    /lab_host="DH10B (phage-resistant)"
                    /note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a
                    modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
                    strand cDNA was primed with a Not I - oligo(dT) primer [5',
                    TGTTACCAATCTGAATGGAGCGCGCTAGCTTTTCTTTTCTTTT 3'],
                    double-stranded cDNA was ligated to Eco RI adaptors
                    (Pharmacia), digested with Not I and cloned into the Not I
                    and Eco RI sites of the modified pT73 vector. Library
                    went through one round of normalization. Tissue samples
                    provided by Dr. Brian Dieckgraefe (Washington University,
                    dieck@m.wustl.edu); colonic mucosa represents a range of
                    disease involvement from mild cryptitis to severe
                    ulceration, fibrosis, and degeneration. Library
                    constructed by Bento Soares and M. Fatima Bonaldo. "
```

```

ORIGIN
Query Match      29.3%; Score 441; DB 9; Length 481;
Best Local Similarity 100.0%; Pred. No. 1.5e-139;
Matches 441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1055 TGCTTCAGTTGTAATGGAATAATGGCTGGGAAAGAGGTGGTGTCCTTAATTGTTT 1114
DB 447 TGCTTCAGTTGTAATGGAATAATGGCTGGGAAAGAGGTGGTGTCCTTAATTGTTT 388
QY 1115 TACTTGTAACTCTCTTGTGCCCCCTGGGCACTTGGCCCTTGTCTCTCAGTGCTT 1174
DB 387 TACTTGTAACTCTCTTGTGCCCCCTGGGCACTTGGCCCTTGTCTCTCAGTGCTT 328
QY 1175 CCCTTTGACATGGGAAGAGTGTGGCCAAATCCCATCTTCTTGCACTCAACGCT 1234
DB 327 CCCTTTGACATGGGAAGAGTGTGGCCAAATCCCATCTTCTTGCACTCAACGCT 268
QY 1235 GTGGCTCAGGGCTGGGTGGCAGAGGAGGCTTTCACCTTATATCTGTGTATCCAG 1294
DB 267 GTGGCTCAGGGCTGGGTGGCAGAGGAGGCTTTCACCTTATATCTGTGTATCCAG 208
QY 1295 GGCTCCAGACTTCTCTCTGCTGCCCCCACTGCACCTCTCCCTTATCTCCTT 1354
DB 207 GGCTCCAGACTTCTCTCTGCTGCCCCCACTGCACCTCTCCCTTATCTCCTT 148
QY 1355 CTGGCTCCCCAGCCCACTTGTGCTTCTTGTGCCCCCTGGGTCACTCCCTCACTCTG 1414
DB 147 CTGGCTCCCCAGCCCACTTGTGCTTCTTGTGCCCCCTGGGTCACTCCCTCACTCTG 88
QY 1415 ACTCTGACTATGGCAGCAGAACACAGGGGCTTGCCCCAGTGGATTTTCATGGTGCATTA 1474
```

```

DB 87 ACTCTGACTATGGCAGCAGAACACCGGCTGGCCAGTGGATTTTCATGGTGCATTA 28
QY 1475 AAAAGAGAAAATCGCAACCAA 1495
DB 27 AAAAGAGAAAATCGCAACCAA 7

RESULT 36
LOCUS   AI803690
DEFINITION tc19e02.x1 Soares NhHMPu_S1 Homo sapiens cDNA clone IMAGE:2064314
            3', mRNA sequence.
ACCESSION AI803690
VERSION   AI803690.1  GI:5369162
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 476)
            NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE     Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            This clone is available royalty-free through LNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 679 Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 453.
FEATURES    source
            1..476
                Location/Qualifiers
                1..476
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:2064314"
                    /tissue_type="Pooled human melanocyte, fetal heart, and
                    pregnant uterus"
                    /lab_host="DH10B"
                    /clone_lib="Soares NhHMPu S1"
                    /note="Organ: mixed (see below); Vector: pT73D-Pac
                    (Pharmacia) with a modified polylinker; Site 1: Not I;
                    Site 2: Eco RI; Equal amounts of plasmid DNA from three
                    normalized libraries (melanocyte 2NbHM, pregnant uterus
                    NbHPU, and fetal heart NbHH19W) were mixed, and ss circles
                    were made in vitro. Following HAP purification, this DNA
                    was used as tracer in a subtractive hybridization
                    reaction. The driver was PCR-amplified cDNAs from pools of
                    5,000 clones made from the same 3 libraries. The pools
                    consisted of I.M.A.G.E. clones 260232-265223,
                    340488-345479, and 484488-489479."
ORIGIN
Query Match      29.1%; Score 438; DB 9; Length 476;
Best Local Similarity 100.0%; Pred. No. 1.6e-138;
Matches 438; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1052 AATTGCTTCAGTTGTAATGGAATAATGGCTGGGAAAGAGGTGGTGTCCTTAATTG 1111
DB 438 AATTGCTTCAGTTGTAATGGAATAATGGCTGGGAAAGAGGTGGTGTCCTTAATTG 379
QY 1112 TTTTACTTGTAACTCTCTTGTGCCCCCTGGGCACTTGGCCCTTGTCTCTCAGTG 1171
DB 378 TTTTACTTGTAACTCTCTTGTGCCCCCTGGGCACTTGGCCCTTGTCTCTCAGTG 319
QY 1172 CTTCCCTTTGACATGGGAAGAGTGTGGCCAAATCCCATCTTCTTGCACTCAACG 1231
DB 318 CTTCCCTTTGACATGGGAAGAGTGTGGCCAAATCCCATCTTCTTGCACTCAACG 259
QY 1232 TCCTGTGGCTCAGGGCTGGGTGGCAGAGGAGGCTTACCTTATATCTGTGTGTATC 1291
```


Db	258	TCTGTGGCTCAGGCTGGGGTGGCAGAGGAGGCCCTTCACTTATATCTGTTGTTATC	199
Qy	1292	CAGGGCTCCAGACTTCTCTCTGCTGCCCTGCCCACTGCACCCCTCTCCCTTATCTATCTC	1351
Db	198	CAGGGCTCCAGACTTCTCTCTGCTGCCCTGCCCACTGCACCCCTCTCCCTTATCTATCTC	139
Qy	1352	CTTCTCGGCTCCAGCCAGCTTGGCTTCTTGTCCCTCTGCTGGGTCACTCCCTCCACT	1411
Db	138	CTTCTCGGCTCCAGCCAGCTTGGCTTCTTGTCCCTCTGCTGGGTCACTCCCTCCACT	79
Qy	1412	CTGACTCTGACTATGGCAGCAGAACACACAGGGCCTGGCCAGTGGATTTCATGGTATCA	1471
Db	78	CTGACTCTGACTATGGCAGCAGAACACACAGGGCCTGGCCAGTGGATTTCATGGTATCA	19
Qy	1472	TTAAAAAGAAAAATCGC	1489
Db	18	TTAAAAAGAAAAATCGC	1
RESULT 37			
LOCUS	BM712008		
DEFINITION	UI-E-EJ0-abf-o-21-0-UI.r1 UI-E-EJ0 Homo sapiens cDNA clone		
ACCESSION	BM712008		
VERSION	1		
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Bonaldo,M.P., Lennon,G. and Soares,M.B.		
TITLE	Normalization and subtraction: two approaches to facilitate gene discovery		
JOURNAL	Genome Res. 6 (9), 791-806 (1996)		
MEDLINE	97044477		
PUBMED	8889548		
COMMENT	Contact: Soares, MB Coordinated Laboratory for Computational Genomics University of Iowa 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA Tel: 319 335 8250 Fax: 319 335 9565 Email: bento-soares@uiowa.edu Tissue Procurement: Dr. Gregg Hageman cDNA Library preparation: Dr. M. Bento Soares, University of Iowa cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com). Seq primer: M13 Reverse.		
FEATURES	Location/Qualifiers		
source	1..474		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="UI-E-EJ0-abf-o-21-0-UI"		
	/tissue_type="fetal eyes, lens, eye anterior segment, optic nerve, retina, Retina Foveal and Macular, RPE and Choroid"		
	/dev_stage="fetal and adult"		
	/lab_host="DH10B (Life Technologies) (T1 phage resistant)"		
	/clone_lib="UI-E-EJ0"		
	/notes="Organ: eye; Vector: p77T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-EJ0 is a subtracted cDNA library constructed according to Bonaldo, Lennon and Soares. Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p77T3-Pac vector. The oligonucleotide used to prime the synthesis of		
first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are: fetal eyes, AGAATCAGA; lens, CGATTAGGA; eye anterior segment, AATGCCCAT; optic nerve, CCATTAGTG; retina, CCGCG; Retina Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."			
ORIGIN			
Query Match	29.0%;	Score 437;	DB 12; Length 474;
Best Local Similarity	100.0%;	Pred. No. 3.5e-138;	
Matches 437;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	951	GGATTTTAGGTTGATTTCTGGATCAGGGATACCACTTCTCTGCCACACCCCGACGAC	1010
Db	1	GGATTTTAGGTTGATTTCTGGATCAGGGATACCACTTCTCTGCCACACCCCGACGAC	60
Qy	1011	GGGCTAGAAAAATTTGTTGAGATTTTATATCATCTTGTCAAAATGCTTCAGTTGTAAT	1070
Db	61	GGGCTAGAAAAATTTGTTGAGATTTTATATCATCTTGTCAAAATGCTTCAGTTGTAAT	120
Qy	1071	GTGAAAAATGGCTGGGGAAAGAGGTGCTCCTTAATGTTTACTTGTAACTTGT	1130
Db	121	GTGAAAAATGGCTGGGGAAAGAGGTGCTCCTTAATGTTTACTTGTAACTTGT	180
Qy	1131	CTTGTGCCCCCTGGGCACCTTGGCTTGTCTCTCAGTGTCTTCCCTTTGACATGGAA	1190
Db	181	CTTGTGCCCCCTGGGCACCTTGGCTTGTCTCTCAGTGTCTTCCCTTTGACATGGAA	240
Qy	1191	AGGAGTTTGGCCAAAATCCCATCTTCTTGCACCTCAACGCTGTGGCTCAGGGTGG	1250
Db	241	AGGAGTTTGGCCAAAATCCCATCTTCTTGCACCTCAACGCTGTGGCTCAGGGTGG	300
Qy	1251	GTGGCAGAGGAGGCGCTTACCTTATATCTGTGTTTATCCAGGCTCCAGACTTCCTC	1310
Db	301	GTGGCAGAGGAGGCGCTTACCTTATATCTGTGTTTATCCAGGCTCCAGACTTCCTC	360
Qy	1311	CTCTGCCCTGCCCTGCACCTCTCCCTTATCTATCTCTCTCGGCTCCCGAGCCC	1370
Db	361	CTCTGCCCTGCCCTGCACCTCTCCCTTATCTATCTCTCTCGGCTCCCGAGCCC	420
Qy	1371	AGTCTTGGCTTCTTGTCTC	1387
Db	421	AGTCTTGGCTTCTTGTCTC	437
RESULT 38			
LOCUS	AI129923/c		
DEFINITION	qc41d08.x1 Soares pregnant uterus_NbHPU Homo sapiens cDNA clone		
ACCESSION	AI129923		
VERSION	AI129923.1		
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.		
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index		
JOURNAL	Unpublished (1997)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cyabps-remail.nih.gov This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Insert Length: 670 Std Error: 0.00 Seq primer: -40m13 fwd. RT from Amersham High quality sequence stop: 392. Location/Qualifiers 1..435		


```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1712175"
/sex="female"
/dev stage="adult"
/lab_host="DH10B"
/clone_lib="Soares_pregnant_uterus_NbHPU"
/site="Organ: uterus; Vector: pF773-Pac; Site1: Not I ;
Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5',
AACTGGCAAGTTCGGCGCGCTTTTITTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pF773 vector. Library
went through one round of normalization. Library
constructed by M. Fatima Ronaldo."

```

ORIGIN

Query Match	28.9%;	Score 435;	DB 9;	Length 435;
Best Local Similarity	100.0%;	Pred. No. 1.7e-137;		
Matches 435;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
1055	TGCTTCAGTGTGTAATGTGAAAAATGGCTGGGGAAAGGAGGTGGTCCCTCAATTGTTTT	1114		
435	TGCTTCAGTGTGTAATGTGAAAAATGGCTGGGGAAAGGAGGTGGTCCCTCAATTGTTTT	376		
1115	TACTTGTTAACTTGTTCCTTGTGCCCCCTGGGCACCTTGGCCCTTGTCTGCTCTCAGTGTCTT	1174		
375	TACTTGTTAACTTGTTCCTTGTGCCCCCTGGGCACCTTGGCCCTTGTCTGCTCTCAGTGTCTT	316		
1175	CCCTTTGACATGGGAAGGAGTTGTGCCCAAAATCCCATCTCTCTTGGACCTCAACGTCT	1234		
315	CCCTTTGACATGGGAAGGAGTTGTGCCCAAAATCCCATCTCTCTTGGACCTCAACGTCT	256		
1235	GTGGCTCAGGGCTGGGGTGGCAGAGGAGGCCCTTACCTTATATCTGTGTGTTATCCAG	1294		
255	GTGGCTCAGGGCTGGGGTGGCAGAGGAGGCCCTTACCTTATATCTGTGTGTTATCCAG	196		
1295	GGCTCCAGACTTCTCTCTGCTGCCCTCCACACTGCACCCCTCTCCCTTATCTATCTCTT	1354		
195	GGCTCCAGACTTCTCTCTGCTGCCCTCCACACTGCACCCCTCTCCCTTATCTATCTCTT	136		
1355	CTGGCTCCCCAGCCCAAGTCTTGGCTTCTTGTCCTCTCTGGGGTCAATCCCTCCACTCTG	1414		
135	CTGGCTCCCCAGCCCAAGTCTTGGCTTCTTGTCCTCTCTGGGGTCAATCCCTCCACTCTG	76		
1415	ACTCTGACTATGGCAGCAGAAACACAGGCGCTGGCCCCAGTGGATTCATGGTGATCATTA	1474		
75	ACTCTGACTATGGCAGCAGAAACACAGGCGCTGGCCCCAGTGGATTCATGGTGATCATTA	16		
1475	AAAAAGAAAAATCGC	1489		
15	AAAAAGAAAAATCGC	1		

RESULT 39
BI765897 LOCUS
DEFINITION
 linear mRNA 804 bp EST 25-SEP-2001
6030461646F1.NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5186459 5',
mRNA sequence.
BI765897
BI765897 GI:15757475
EST.
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini, Hominidae; Homo.
1 (bases 1 to 804)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

Email: cgapbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MSC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM11465 row: 1 column: 12
High quality sequence stop: 748.

FEATURES

source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5186459"
/lab_host="DH10B"
/clone.lib="NIH MGC 116"
/note="Organ: pooled colon, kidney, stomach; Vector:
PCMV-SPORT; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2
colons, 62 yo male and 70 yo female. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
023. Note: this is NIH MGC Library."

```

ORIGIN

Query Match	28.8%;	Score 433;	DB 12;	Length 804;
Best Local Similarity	99.6%;	Pred. No. 6.3e-137;		
Matches 533;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
QY	1	CGCGGATCGGACCCAGCAGGTCGGCGGCGGCGGAGGAGGCGGCGGCTCAGCTCC	60	
Db	83	CGCGGATCGGACCCAGCAGGTCGGCGGCGGCGGAGGAGGCGGCGGCTCAGCTCC	142	
QY	61	TCGACCCCGCTGTCGGGCTAGTCCAGCAGGCGGAGCGGCGGCTGGGCGCCATGGCCAG	120	
Db	143	TCGACTCCCGTGTGGGCTAGTCCAGCAGGCGGAGCGGCGGCTGGGCGCCATGGCCAG	202	
QY	121	CCGGCATGAGCGGTGGCGCGACCGGCTGGCGCTGTGTGACGGGGGCGCTCGGGGGCAATC	180	
Db	203	CCCGCATGAGCGGTGGCGCGACCGGCTGGCGCTGTGTGACGGGGGCGCTCGGGGGGCATC	262	
QY	181	GGCGGGCGGCTGGCGCGGCGGCGGCTGGTCAGACGAGGACTGAAGTGTGGGCTGGCGCGC	240	
Db	263	GGCGGGCGGCTGGCGCGGCGGCGGCTGGTCAGACGAGGACTGAAGTGTGGGCTGGCGCGC	322	
QY	241	ACTGTGGGGAACATCGAGAGTGCTGCTGTAATCTAAGAGTGCAGGCTACCCCGGGACT	300	
Db	323	ACTGTGGGGAACATCGAGAGTGCTGCTGTAATCTAAGAGTGCAGGCTACCCCGGGACT	382	
QY	301	TTGATCCCGCTACAGATGTCACTCAAAATGAGAGGACATCCTCTCAATGTTCTCAGGT	360	
Db	383	TTGATCCCGCTACAGATGTCACTCAAAATGAGAGGACATCCTCTCAATGTTCTCAGGT	442	
QY	361	ATCCGTTTCTCAGCACAGCGGTGTAGACATCTGCATCAACAATGCTGGCTTGGCCCGGCT	420	
Db	443	ATCCGTTTCTCAGCACAGCGGTGTAGACATCTGCATCAACAATGCTGGCTTGGACCGGCT	502	
QY	421	GACACCTTGCTCTCAGGAGCACGAGTGTTGGAGGACATGTTCAATGTGAACGTGCTG	480	
Db	503	GACACCTTGCTCTCAGGAGCACGAGTGTTGGAGGACATGTTCAATGTGAACGTGCTG	562	
QY	481	GCCCTCAGCATCTGCACACGGGAGCCCTACAGTCCATGAAGGACCGGAATGTGG	535	
Db	563	GCCCTCAGCATCTGCACACGGGAGCCCTACAGTCCATGAAGGACCGGAATGTGG	617	

RESULT 40

```

AA448177/c
LOCUS      zw83b12.s1 Soares testis_NHT Homo sapiens cDNA clone IMAGE:782783
DEFINITION 3', mRNA sequence.
ACCESSION  AA448177
VERSION     AA448177.1 GI:2161847
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 430)
AUTHORS     Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisler, G., Jost, S.,
            Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J.,
            Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,
            White, Y., Wylie, T., Waterston, R. and Wilson, R.
            WashU-Merck EST Project 1997
            Unpublished (1997)
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LNL; contact the
            IMAGE Consortium (info@image.lnl.gov) for further information.
            Seq primer: -41ml3 fwd. ET from Amersham
            High quality sequence stop: 408.
FEATURES   Location/Qualifiers
            1..430
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:782783"
                /sex="male"
                /lab_host="DH10B"
                /clone_lib="Soares testis NHT"
                /note="Vector: pT73D-Pac (Pharmacia) with a modified
                polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
                was prepared from mRNA obtained from Clontech
                Laboratories, Inc., and primed with a Not I - oligo(dT)
                primer [5].
                TGTACCAATCTGAAGTGGGAGCGCGCCCAATTTTTTTTTTTT 3'.
            Double-stranded cDNA was ligated to Eco RI adaptors
            (Pharmacia), digested with Not I and cloned into the Not I
            and Eco RI sites of the modified pT73 vector. Library
            went through one round of normalization to Cot5, and was
            constructed by Bento Soares and M. Fatima Bonaudo.
ORIGIN
Query Match      28.4%; Score 428; DB 9; Length 430;
Best Local Similarity 100.0%; Pred. No. 4.2e-135;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1064 TGTAAATGTGAAATGGCTGGGAAAGAGAGTGTGTCCCTAATTTGTTTACTTGTTA 1123
DB 430 TGTAAATGTGAAATGGCTGGGAAAGAGAGTGTGTCCCTAATTTGTTTACTTGTTA 371
QY 1124 ACTTGTCTTGTGCCCCCTGGGCACATTGGCCTTTGTCTGTCTCAGTGTCTCCCTTTGAC 1183
DB 370 ACTTGTCTTGTGCCCCCTGGGCACATTGGCCTTTGTCTGTCTCAGTGTCTCCCTTTGAC 311
QY 1184 ATGGGAAGAGTTGTGGCCAAATCCCATCTTTGACCTCACTCACTGTCTGGCTCAG 1243
DB 310 ATGGGAAGAGTTGTGGCCAAATCCCATCTTTGACCTCACTCACTGTCTGGCTCAG 251
QY 1244 GCTGGGGTGGCAGAGGAGGCTTCACTTATATCTGTGTGTATCCAGGGCTCCAG 1303
DB 250 GCTGGGGTGGCAGAGGAGGCTTCACTTATATCTGTGTGTATCCAGGGCTCCAG 191
QY 1304 CTTCCCTCTCTGCTGCCCCACTGCACCTCTCCCTTATCTATCTCTTCCTCGCTCC 1363
DB 190 CTTCCCTCTCTGCTGCCCCACTGCACCTCTCCCTTATCTATCTCTCTTCCTCGCTCC 131

```

```

QY 1364 CCAGCCAGCTTGGCTTTCTTGTCCCTCCTGGGGTTCATCCCTCCTCCTGACTCTGACT 1423
DB 130 CCAGCCAGCTTGGCTTTCTTGTTCCTTGTCCCTCCTGGGGTTCATCCCTCCTGACTCTGACT 71
QY 1424 ATGCACAGACAACACACAGGCGCTGGCCAGTGGATTTTCATGGTGATCATTAATAAGAAA 1483
DB 70 ATGCACAGACAACACACAGGCGCTGGCCAGTGGATTTTCATGGTGATCATTAATAAGAAA 11
QY 1484 AATGCGAA 1491
DB 10 AATGCGAA 3

RESULT 41
CB851997/c
LOCUS      CB851997
DEFINITION UI-CF-FNO-aes-e-22-0-UI.s1 UI-CF-FNO Homo sapiens cDNA clone
            UI-CF-FNO-aes-e-22-0-UI 3', mRNA sequence.
ACCESSION  CB851997
VERSION     CB851997.1 GI:30046840
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 669)
AUTHORS     Bonaudo, M.F., Lennon, G. and Soares, M.B.
            Normalization and subtraction: two approaches to facilitate gene
            discovery
            Genome Res. 6 (9), 791-806 (1996)
            97044477
            MEDLINE
            PUBMED
            8859548
            COMMENT
                McCray Lab
                University of Iowa
                2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
                Tel: 319 356 4866
                Fax: 319 356 7171
                Email: paul-mccray@uiowa.edu
                Tissue Procurement: Dr. M. J. Welsh, University of Iowa
                cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
                cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
                DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
                Clone Distribution: Researchers may obtain clones from Research
                Genetics (www.reagen.com).
            The following repetitive elements were found in this cDNA
            sequence: 1-38, >POLY_A#Simple_repeat (matched complement)
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES   Location/Qualifiers
            1..669
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="UI-CF-FNO-aes-e-22-0-UI"
                /tissue_type="Human Lung Epithelial cells"
                /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                /clone_lib="UI-CF-FNO"
                /note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
                modified polylinker; Site 1: EcoR I; Site 2: Not I;
                UI-CF-FNO is a subtracted cDNA library derived from two
                normalized Human lung epithelial cell libraries (EN1 and
                DUL) The library was subtracted according to according to
                Bonaudo, Lennon and Soares, Genome Research, 6:791-806,
                1996. For additional information, contact:
                bento-soares@uiowa.edu
                TAG_SEQ=None found"
ORIGIN
Query Match      28.4%; Score 427; DB 14; Length 669;
Best Local Similarity 99.7%; Pred. No. 7.4e-135;
Matches 597; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

```

QY 907 CTCCTCTTCCCTCCCAACCTTCATGGCTTGCTCTGCTCTGGATTTAGGCTTGA 966
Db 604 CTCCTCTTCCCTCCCAACCTTCATGGCTTGCTCTGCTCTGGATTTAGGCTTGA 545
QY 967 TTCTGTGATCAGGGATACACTTCTCTGCTCCACCCGACACAGGGGCTAGAAAATTCT 1026
Db 544 TTCTGTGATCAGGGATACACTTCTCTGCTCCACCCGACACAGGGGCTAGAAAATTCT 485
QY 1027 TTGAGATTTTATATCATCTTGTCAAAATGCTTCAGTCTTAATGTGAAAAATGGCTGG 1086
Db 484 TTGAGATTTTATATCATCTTGTCAAAATGCTTCAGTCTTAATGTGAAAAATGGCTGG 425
QY 1087 GGAAGAGAGGTGGTGTCCCTAAATGTTTACTTGTAACTTGTCTTGTGCCCCCTGGGCA 1146
Db 424 GGAAGAGAGGTGGTGTCCCTAAATGTTTACTTGTAACTTGTCTTGTGCCCCCTGGGCA 365
QY 1147 CTTGGCCCTTGTCTGCTCTCACTGCTTCCCTTTGACATGGGAAAGGATTTGCGCCAA 1206
Db 364 CTTGGCCCTTGTCTGCTCTCACTGCTTCCCTTTGACATGGGAAAGGATTTGCGCCAA 306
QY 1207 ATCCCATCTTCTTGCACTCAACGCTCTGTGCTCAGGGCTGGGTGGCAGAGGAGGCC 1266
Db 305 ATCCCATCTTCTTGCACTCAACGCTCTGTGCTCAGGGCTGGGTGGCAGAGGAGGCC 246
QY 1267 TTCACCTTATATCTGTGTGTATCCAGGGCTCCAGACTTCTCTCTGCTGCCCCCACT 1326
Db 245 TTCACCTTATATCTGTGTGTATCCAGGGCTCCAGACTTCTCTCTGCTGCCCCCACT 186
QY 1327 GCACCCCTCCCTTATCTATCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1386
Db 185 GCACCCCTCCCTTATCTATCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 126
QY 1387 CCCCTCTCGGGCTCATCCCTCCACTCTGACTCTGACTATGCTGCTGCTGCTGCTGCTGCT 1446
Db 125 CCCCTCTCGGGCTCATCCCTCCACTCTGACTCTGACTATGCTGCTGCTGCTGCTGCTGCT 66
QY 1447 GCGCCAGTGGATTTAATGGTATCATTAAGAAAGAAATGCGCAACCAAAAAA 1505
Db 65 GCGCCAGTGGATTTAATGGTATCATTAAGAAAGAAATGCGCAACCAAAAAA 7

RESULT 42
BE796469
LOCUS
DEFINITION
601589817f1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3943948 5',
mRNA sequence.
ACCESSION
BE796469
VERSION
BE796469.1 GI:10217667
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://img.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Cloning Distribution by: Incyte Genomics, Inc.
Found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLC#800 row: a column: 05
High quality sequence stop: 782.
Location/Qualifiers
1. .1063
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
FEATURES
source

```

```

/clone="IMAGE:3943948"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_7"
/notes="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

Query Match 28.2%; Score 425; DB 10; Length 1063;
Best Local Similarity 100.0%; Pred. No. 2.9e-134; Mismatches 0; Indels 0; Gaps 0;
Matches 425; Conservative 0;

QY	142	GACCGCTGGCGCTGGTCAACGGGGCCCTCGGGGGGCATCGGGCGCGCGTGGCCCGGGCC	201
Db	166	GACCGCTGGCGCTGGTCAACGGGGCCCTCGGGGGGCATCGGGCGCGCGTGGCCCGGGCC	225
QY	202	CTGCTCCAGCAGGGGACTGAAGGTGGTGGGCTGGCCCGCAGCTGGGGCAACATCGAGGAG	261
Db	226	CTGCTCCAGCAGGGGACTGAAGGTGGTGGGCTGGCCCGCAGCTGGGGCAACATCGAGGAG	285
QY	262	CTGGCTGCTGAATGAAGAGTGCAGGCTACCCCGGGACTTTGATCCCTACAGATGTGAC	321
Db	286	CTGGCTGCTGAATGAAGAGTGCAGGCTACCCCGGGACTTTGATCCCTACAGATGTGAC	345
QY	322	CTATCAATGAAGAGGACATCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGT	381
Db	346	CTATCAATGAAGAGGACATCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGT	405
QY	382	GTAGACATCTGCATCAACATGCTGGTGGCCCGGCTGACACCTGCTCTCAGGCGAGC	441
Db	406	GTAGACATCTGCATCAACATGCTGGTGGCCCGGCTGACACCTGCTCTCAGGCGAGC	465
QY	442	ACCACTGTTGGAGGACATGTTCAATGTGAACGTTGTCGCTGGCCCTCAGCATCTGCACACGG	501
Db	466	ACCACTGTTGGAGGACATGTTCAATGTGAACGTTGTCGCTGGCCCTCAGCATCTGCACACGG	525
QY	502	GAGCCTTACAGTCCATGAAGAGGCGGAATGTGACGATGGGACATCATTAACATCAAT	561
Db	526	GAGCCTTACAGTCCATGAAGAGGCGGAATGTGACGATGGGACATCATTAACATCAAT	585
QY	562	AGCAT 566	
Db	586	AGCAT 590	

RESULT 43
AW452081/c
LOCUS
DEFINITION
UI-H-B13-ahn-c-06-0-UI.sl NCI_CGAP_Sub5 Homo sapiens cDNA clone
IMAGE:2737163 3', mRNA sequence.
ACCESSION
AW452081
VERSION
AW452081.1 GI:6992857
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 444)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first

strand cDNA and therefore this may represent a bonafide poly A
tail. cDNA library Preparation: M.B. Soares Lab Clone distribution:
NCI-CGAP clone distribution information can be found through the
I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: M13 Forward
POLYA=Yes.

FEATURES
source
1. .444
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2737163"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI_CGAP_Sub5"
/note="Vector: p773D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; NCI CGAP Sub5
is a subtracted library derived from NCI CGAP Sub4. The
NCI CGAP Sub5 library had 3 million recombinants. A
single-stranded DNA preparation of NCI_CGAP_Sub4 was used
as a tracer in a subtractive hybridization with a driver
comprising: the IMAGE pool (NCI CGAP Kid3 pool 1 LLAM
3334-3337, 3682-3683, 3798-3803 (IMAGE Clones)
132376-132391, 1456008-1456775, 1500552-1502855);
NCI CGAP Kid5 pool 1 LLAM 3338-3342, 3722-3725, 3776-3778
(IMAGE Clones) 1323912-1325831, 1473368-1472903,
1492104-1493255); NCI CGAP Lu5 pool 1 LLAM 3575-3582,
3851-3854 (IMAGE Clones)
1414920-1417991, 1520904-1522439); NCI CGAP GC4 pool 1 LLAM
3164-3167, 3716-3720, 3733-3735 (IMAGE Clones)
1257096-1258631, 1469064-1470983, 1475592-1476743);
NCI CGAP Pr22 pool 1 LLAM 2457-2459, 2758-2759, 3062-3068
(IMAGE Clones) 985608-986759, 1101192-1101959,
1217928-1220615); NCI CGAP Co10 pool 1 LLAM 2644-2653,
2871-2872 (IMAGE Clones)
1057416-1061255, 1144584-1145351). (10% of the driver
population), plus a pool of 3,840 arrayed clones from
NCI CGAP Sub1 (IMAGE Clones) 2708616-2710535) and
NCI CGAP Sub2 (IMAGE Clones) 2710536-2712455) (10% of
the driver population), plus a pool of 11,136 clones from
NCI CGAP Sub3 (IMAGE Clones) 2712456-2723591) (10% of the
driver population), plus a pool of 5,472 clones from
NCI CGAP Sub4 (IMAGE Clones) 2723592-2728969) (70% of the
driver population). Subtraction was performed as
previously described (Bonaldi, Lennon & Soares (1996):
Normalization and Subtraction: Two Approaches To
Facilitate Gene Discovery. Genome Research 6, 791-806.
TAG_TISSUE=prostate
TAG_LIB=NCI_CGAP_Pr22
TAG_SEQ=AAGTG"

ORIGIN
Query Match 28.2%; Score 424; DB 10; Length 444;
Best Local Similarity 100.0%; Pred. No. 9, 3e-134;
Matches 42%; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1067 AAATGTGAAATATGGCTGGGAAAGAGGTGGTGTCCCTAATGTTTACTGTAACT 1126
Db 444 AAATGTGAAATATGGCTGGGAAAGAGGTGGTGTCCCTAATGTTTACTGTAACT 385
QY 1127 TGTTCCTGTGGCCCTGGGCATGGCCTTGTCTGTCTCAGTGTCTTCCCTTTCATG 1186
Db 384 TGTTCCTGTGGCCCTGGGCATGGCCTTGTCTGTCTCAGTGTCTTCCCTTTCATG 325
QY 1187 GGAAAGAGTGTGGCAAAATCCCATCTTCTTGACCTCAACGCTGTGGCTCAGGC 1246
Db 324 GGAAAGAGTGTGGCAAAATCCCATCTTCTTGACCTCAACGCTGTGGCTCAGGC 265
QY 1247 TGGGGTGGCAGAGGAGCCCTTACCTTATCTGTGTTATCCAGGCTCCAGACTT 1306
Db 264 TGGGGTGGCAGAGGAGCCCTTACCTTATCTGTGTTATCCAGGCTCCAGACTT 205
QY 1307 CCTCCCTGCTGCCCCACTGCACCTCTCCCCCTTATCTATCTCTCTTCGCGTCCCCA 1366

Db 204 CTTCTCTGCTGCTGCCACTGCACCTCTCCCTTATCTATCTCTCTCGGCTCCCCA 145
QY 1367 GCCAGTCTTGGCTTCTTGTCCCTCTCTGGGTCATCCCTCCACTCTGACTCTGACTATG 1426
Db 144 GCCAGTCTTGGCTTCTTGTCCCTCTCTGGGTCATCCCTCCACTCTGACTCTGACTATG 85
QY 1427 GCAGCAGAACACAGGCGCTGGGCCAGTGGATTTTCATGTCATTTAAAAAGAAAAAT 1486
Db 84 GCAGCAGAACACAGGCGCTGGGCCAGTGGATTTTCATGTCATTTAAAAAGAAAAAT 25
QY 1487 CGCA 1490
Db 24 CGCA 21

RESULT 44
BE869917
LOCUS
DEFINITION
601446563F1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3850574 5',
mRNA sequence.
614 bp mRNA linear EST 20-OCT-2000
BE869917
VERSION
BE869917.1 GI:10318693
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (Bases 1 to 614)
NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE
Unpublished (1999)
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9569 row: n column: 15
High quality sequence stop: 602.
Location/Qualifiers
1. 614
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3850574"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_65"
/note="Organ: Colon; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.8 Kb. Library constructed by Life
Technologies."

ORIGIN
Query Match 28.1%; Score 423; DB 10; Length 614;
Best Local Similarity 99.8%; Pred. No. 1.8e-133;
Matches 54%; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 33 GCAGAGAGAGCGCGCGCGTCCAGTCTCTCGACCCCGCTGTGGGTAGTCCAGCAGGC 92
Db 1 GCAGAGAGAGCGCGCGCGTCCAGTCTCTCGACCCCGTGTGGGTAGTCCAGCAGGC 60
QY 93 GGAACGCGCGCGTGTGGGCCCATGGCCAGGCCCGCATGGCGCGCGCGCGCTGCG 152
Db 61 GGACGCGCGCGTGTGGGCCCATGGCCAGGCCCGCATGGCGCGCGCGCGCTGCG 120
QY 153 GCTGTGTGACGGGGCGCTCGGGGGCGCATCGCGCGCGCGCGCTGGCCCCCGGCTGTCAGCA 212
Db 121 GCTGTGTGACGGGGCGCTCGGGGGCGCATCGCGCGCGCGCGCTGGCCCCCGGCTGTCAGCA 179


```
source
1..751
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4374184"
/lab_host="DH10B"
/clone_lib="NIH_MGC_97"
/notes="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI
(gtcgag); Oligo-dr primed using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average
insert size 2.2 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIMH/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC Library."

ORIGIN
Query Match 27.8%; Score 418; DB 10; Length 751;
Best Local Similarity 100.0%; Pred. No. 8e-132;
Matches 418; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGATCGGACCAAGCAGGTTCGGCGCGCGGAGAGAGCGCGCGGTTCAGTCC 60
DB 107 CGCGATCGGACCAAGCAGGTTCGGCGCGCGGAGAGAGCGCGCGGTTCAGTCC 166

QY 61 TCGACCCCGGTGTTCGGGTAGTCCAGCGAGCGGCGCGGTTCAGTCCAGTCC 120
DB 167 TCGACCCCGGTGTTCGGGTAGTCCAGCGAGCGGCGCGGTTCAGTCCAGTCC 226

QY 121 CCCGGCATGAGAGCGGTTCGGGTAGTCCAGCGAGCGGCGCGGTTCAGTCCAGTCC 180
DB 227 CCCGGCATGAGAGCGGTTCGGGTAGTCCAGCGAGCGGCGCGGTTCAGTCCAGTCC 286

QY 181 GCGCGCGCGGTTCGGGTAGTCCAGCGAGCGGCGCGGTTCAGTCCAGTCCAGTCC 240
DB 287 GCGCGCGCGGTTCGGGTAGTCCAGCGAGCGGCGCGGTTCAGTCCAGTCCAGTCC 346

QY 241 ACTGTGGGCAACATCGAGAGTGGCTGCTGAATGAAGTGAAGTGAAGTGAAGTGAAGT 300
DB 347 ACTGTGGGCAACATCGAGAGTGGCTGCTGAATGAAGTGAAGTGAAGTGAAGTGAAGT 406

QY 301 TTGATCCCTACAGATGACCTATCAATGAAGAGGACATCTCCATGTTCTCAGT 360
DB 407 TTGATCCCTACAGATGACCTATCAATGAAGAGGACATCTCCATGTTCTCAGT 466

QY 361 ATCCGTTCTCAGCAGCGGTGTAGACATCTGCATCAACAATGCTGGCTTGGCCCGGC 418
DB 467 ATCCGTTCTCAGCAGCGGTGTAGACATCTGCATCAACAATGCTGGCTTGGCCCGGC 524

RESULT 47
BI464353
LOCUS 603204005F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5269728 5',
DEFINITION mRNA sequence.
ACCESSION BI464353
VERSION 1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 757)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
```

```
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11680 row: j column: 01
High quality sequence stop: 707.

FEATURES
source
1..757
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5269728"
/lab_host="DH10B"
/clone_lib="NIH_MGC_97"
/notes="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI
(gtcgag); Oligo-dr primed using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average
insert size 2.2 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIMH/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC Library."

ORIGIN
Query Match 27.7%; Score 417; DB 12; Length 757;
Best Local Similarity 99.8%; Pred. No. 1.7e-131;
Matches 537; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 CGCGATCGGACCAAGCAGGTTCGGCGCGCGGAGAGAGCGCGCGGTTCAGTCC 60
DB 54 CGCGATCGGACCAAGCAGGTTCGGCGCGCGGAGAGAGCGCGCGGTTCAGTCC 113

QY 61 TCGACCCCGGTGTTCGGGTAGTCCAGCGAGCGGCGCGGTTCAGTCCAGTCC 120
DB 114 TCGACCCCGGTGTTCGGGTAGTCCAGCGAGCGGCGCGGTTCAGTCCAGTCC 173

QY 121 CCCGGCATGAGAGCGGTTCGGGTAGTCCAGCGAGCGGCGCGGTTCAGTCCAGTCC 180
DB 174 CCCGGCATGAGAGCGGTTCGGGTAGTCCAGCGAGCGGCGCGGTTCAGTCCAGTCC 233

QY 181 GCGCGCGCGGTTCGGGTAGTCCAGCGAGCGGCGCGGTTCAGTCCAGTCCAGTCC 240
DB 234 GCGCGCGCGGTTCGGGTAGTCCAGCGAGCGGCGCGGTTCAGTCCAGTCCAGTCC 293

QY 241 ACTGTGGGCAACATCGAGAGTGGCTGCTGAATGAAGTGAAGTGAAGTGAAGTGAAGT 300
DB 294 ACTGTGGGCAACATCGAGAGTGGCTGCTGAATGAAGTGAAGTGAAGTGAAGTGAAGT 353

QY 301 TTGATCCCTACAGATGACCTATCAATGAAGAGGACATCTCCATGTTCTCAGC 359
DB 354 TTGATCCCTACAGATGACCTATCAATGAAGAGGACATCTCCATGTTCTCAGC 413

QY 360 TATCCGTTCTCAGCAGCGGTGTAGACATCTGCATCAACAATGCTGGCTTGGCCCGGC 419
DB 414 TATCCGTTCTCAGCAGCGGTGTAGACATCTGCATCAACAATGCTGGCTTGGCCCGGC 473

QY 420 TGACACCCCTGCTCTCAGCAGCGGTTCGGGTAGGAGGACATGTTCAATGTAAGTGTCT 479
DB 474 TGACACCCCTGCTCTCAGCAGCGGTTCGGGTAGGAGGACATGTTCAATGTAAGTGTCT 533

QY 480 GSCCCTCAGCATCTGCACACGCGGAGCTTACCAGTCCATGAAGGAGCGGAATGTGAC 537
DB 534 GSCCCTCAGCATCTGCACACGCGGAGCTTACCAGTCCATGAAGGAGCGGAATGTGAC 591
```

```
RESULT 48
AL568188/c
LOCUS AL568188 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
DEFINITION CS0DF035YM23 3-PRIME, mRNA sequence.
```


AL568188
 VERSION AL568188.2 GI:31291036
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 980)
 Li.W.B., Gruber.C., Jesse.J. and Pollayes.D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 Contact: Genoscope
 On Feb 16, 2001 this sequence version replaced gi:12922282.
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 was not normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 8845.f For more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CSODF035AG12NPLcluster=8845.f. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CSODF035AG12NPL.
 Location/Qualifiers
 1. 980
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSODF035YM23"
 /tissue_type="FETAL BRAIN"
 /dev_stage="fetal"
 /clone_lib="Homo sapiens FETAL BRAIN"
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double-strand cDNA was digested with Not I and
 cloned into the Not I and EcoRV sites of the pCMVSPORT 6
 vector. Library was not normalized."
 ORIGIN
 Query Match 27.6%; Score 415; DB 9; Length 980;
 Best Local Similarity 100.0%; Pred. No. 7.4e-131;
 Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 925 CCCTTCATGGCTCCCTCCTCGATTTAGGTGTCATTTCTGGATCAGGGATA 984
 Db 569 CCCTTCATGGCTCCCTCCTCGATTTAGGTGTCATTTCTGGATCAGGGATA 510
 QY 985 CCACTTCCTGTCACACCCGACAGGGGCTAGAAAATTTGTTGAGATTTTATATCAT 1044
 Db 509 CCACTTCCTGTCACACCCGACAGGGGCTAGAAAATTTGTTGAGATTTTATATCAT 450
 QY 1045 CTTGTGAAATGCTTCAGTTGTAATGTGAAATAGGGCTGGGAAAGGAGGTGCTGCC 1104
 Db 449 CTTGTGAAATGCTTCAGTTGTAATGTGAAATAGGGCTGGGAAAGGAGGTGCTGCC 390
 QY 1105 CTAATTCCTTTACTTGTAACTTCTCTGTCCTTGGCCCTGGGACATTCGCTGCTC 1164
 Db 389 CTAATTCCTTTACTTGTAACTTCTCTGTCCTTGGGACATTCGCTGCTC 330
 QY 1165 TCAGTGTCTTCCCTTTGACATGGGAAAGGAGTTGTGGCCAAATCCCATCTTCTTGCAC 1224
 Db 329 TCAGTGTCTTCCCTTTGACATGGGAAAGGAGTTGTGGCCAAATCCCATCTTCTTGCAC 270
 QY 1225 CTCACGCTCTGTGGCTCAGGGCTGGGTGGCAGAGGAGGCTTACCTTATATCTGTG 1284
 Db 269 CTCACGCTCTGTGGCTCAGGGCTGGGTGGCAGAGGAGGCTTACCTTATATCTGTG 210
 QY 1285 TGTATATCCAGGCTCCAGACTTCTCTCTGCTGCCCTGCCCTGACACCTCTCCCC 1339
 Db 209 TGTATATCCAGGCTCCAGACTTCTCTCTGCTGCCCTGCCCTGACACCTCTCCCC 155

RESULT 49
 AW050632/c
 LOCUS
 DEFINITION
 ACCESSION
 AW050632
 KEYWORDS
 EST.
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 457)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapsb@mail.nih.gov
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 978 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 457.
 Location/Qualifiers
 1. 457
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2558534"
 /tissue_type="colonic mucosa from 3 patients with Crohn's
 disease"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="Soares Dieckgraebe colon_NHCD"
 /note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a
 modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer [5'
 TGTACCAATCTGAAGTGGGCGGCGCTCTTTTTTTTTTTTTTTT 3']
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization. Tissue samples
 provided by Dr. Brian Dieckgraebe (Washington University,
 dickim.wustl.edu); colonic mucosa represents a range of
 disease involvement from moderate to severe Crohn's
 disease; samples include both perforating (fistulas) and
 non-perforating samples. Library constructed by Bento
 Soares and M. Fatima Bonaldo."

FEATURES
 source

ORIGIN

Query Match 27.0%; Score 406; DB 9; Length 457;
 Best Local Similarity 99.8%; Pred. No. 1.2e-127;
 Matches 456; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1037 TATATCATCTGTGCAAAATTCCTTCAGTTGTAATGTGAAATAGGGCTGGGAAAGAGG 1096
 Db 457 TATATCATCTGTGCAAAATTCCTTCAGTTGTAATGTGAAATAGGGCTGGGAAAGAGG 398
 QY 1097 TGTGTGCCCTAAATGTTTTACTTTGTTAACTTGTTCGTCCTGGGACATTTGGCCTTT 1156
 Db 397 TGTGTGCCCTAAATGTTTTACTTTGTTAACTTGTTCGTCCTGGGACATTTGGCCTTT 338
 QY 1157 GTCTGTCTCAGTGTCTTCCCTTTTGACATGGGAAAGAGTTGTGGCCAAATCCCATCT 1216
 Db 337 GTCTGTCTCAGTGTCTTCCCTTTTGACATGGGAAAGAGTTGTGGCCAAATCCCATCT 278
 QY 1217 TCTTGACCTCAACGCTCTGTGGCTCAGGGCTGGGTGGCAGAGGAGGCTTACCTTAT 1276
 Db 277 TCTTGACCTCAACGCTCTGTGGCTCAGGGCTGGGTGGCAGAGGAGGCTTACCTTAT 218
 QY 1277 ATCTGTGTTGTTATCCAGGCTCCAGACTTCTCTCTGCTGCCCTGCCCCACTGACCTCTC 1336
 Db 217 ATCTGTGTTGTTATCCAGGCTCCAGACTTCTCTCTGCTGCCCTGCCCCACTGACCTCTC 158

Query Match		100.0%;	Score 1505;	DB 6;	Length 1505;
Best Local Similarity		100.0%;	Pred. No. 1.5e-278;	Indels 0;	Gaps 0;
Matches 1505;		Conservative 0;	Mismatches 0;		
Qy	1	CGCGATCGGACCCAAAGCAGGTGCGCGCGCGCGGCGAGAGCGCGCGGGCGTCAAGTCC	60		
Db	1	CGCGATCGGACCCAAAGCAGGTGCGCGCGCGCGGCGAGAGCGCGCGGGCGTCAAGTCC	60		
Qy	61	TCGACCCCGGTGTCGGGTAGTCCAGCGAGCGGAGCGGCGCGCGTGGGCCCATGCCAGG	120		
Db	61	TCGACCCCGGTGTCGGGTAGTCCAGCGAGCGGAGCGGCGCGCGTGGGCCCATGCCAGG	120		
Qy	121	CCGCGCATGAGCGGTGCGCGCGACCGCTGCGGTGCGTGGAGCGGGGCGTCCGGGGGCA	180		
Db	121	CCGCGCATGAGCGGTGCGCGCGACCGCTGCGGTGCGTGGAGCGGGGCGTCCGGGGGCA	180		
Qy	181	GGCGCGGCGGTGGCGCGCGGCGCTGGTCCAGCAGGAGCTGAAGGTGGTGGCGCGGCT	240		
Db	181	GGCGCGGCGGTGGCGCGCGGCGCTGGTCCAGCAGGAGCTGAAGGTGGTGGCGCGGCT	240		
Qy	241	ACTGTGGGCAACATCGAGGAGCTGGTGTGAATGTAAGAGTGCAAGGCTACCCGGGACT	300		
Db	241	ACTGTGGGCAACATCGAGGAGCTGGTGTGAATGTAAGAGTGCAAGGCTACCCGGGACT	300		
Qy	301	TTGATCCCTACAGATGTGACCTATCAAAATGAAGAGGACATCCCTCCATGTTCTCAGCT	360		
Db	301	TTGATCCCTACAGATGTGACCTATCAAAATGAAGAGGACATCCCTCCATGTTCTCAGCT	360		
Qy	361	ATCCGTTCTCAGCAGCGGTGTAGACATCTGCATCAACAATGCTGGCTTGGCCCGGCT	420		
Db	361	ATCCGTTCTCAGCAGCGGTGTAGACATCTGCATCAACAATGCTGGCTTGGCCCGGCT	420		
Qy	421	GACACCTGCTCTCAGGAGACACAGTGGTGGAGGACATGTTCAATGTGAAGTGGCTG	480		
Db	421	GACACCTGCTCTCAGGAGACACAGTGGTGGAGGACATGTTCAATGTGAAGTGGCTG	480		
Qy	481	GCCCTCAGCATCTGCACAGGGAAGCTTACCAGTCCATGAAGGAGCGGAATGCGAGAT	540		
Db	481	GCCCTCAGCATCTGCACAGGGAAGCTTACCAGTCCATGAAGGAGCGGAATGCGAGAT	540		
Qy	541	GGGCACATCAATTAACATCAATAGCATGTCTGCCACCGAGTGTTACCCCTGCTGTGACC	600		
Db	541	GGGCACATCAATTAACATCAATAGCATGTCTGCCACCGAGTGTTACCCCTGCTGTGACC	600		
Qy	601	CACCTTATAGTGCACCAAGTATGCGTCACTGCGCTGACAGAGGAGTCAAGCAAGAG	660		
Db	601	CACCTTATAGTGCACCAAGTATGCGTCACTGCGCTGACAGAGGAGTCAAGCAAGAG	660		
Qy	661	CTTCGGGAGGCCAGACCCACATCCGAGCCACGTGCATCTCTCCAGGTGTGTTGAGACA	720		
Db	661	CTTCGGGAGGCCAGACCCACATCCGAGCCACGTGCATCTCTCCAGGTGTGTTGAGACA	720		
Qy	721	CAATTGCGCTTCAAACTCCAGCAAGGACCCCTGAGAGGAGGAGTGCACCTATGAGCAA	780		
Db	721	CAATTGCGCTTCAAACTCCAGCAAGGACCCCTGAGAGGAGGAGTGCACCTATGAGCAA	780		
Qy	781	ATGAAGTGTCTCAAAACCGAGAGTGTGCCAGGCTGTTATCTAGTCTCAGCACCCCC	840		
Db	781	ATGAAGTGTCTCAAAACCGAGAGTGTGCCAGGCTGTTATCTAGTCTCAGCACCCCC	840		
Qy	841	GCACACATCCAGATTGGAGACATCCAGATGAGGCCACGAGCAGGTGACCTAGTGACTG	900		
Db	841	GCACACATCCAGATTGGAGACATCCAGATGAGGCCACGAGCAGGTGACCTAGTGACTG	900		
Qy	901	TGGGAGCTCCTCTTCCCTCCCGACCTTCATGGCTTGGCTTGCCTCCCTCTGGATTTAGG	960		
Db	901	TGGGAGCTCCTCTTCCCTCCCGACCTTCATGGCTTGGCTTGCCTCCCTCTGGATTTAGG	960		
Qy	961	TGTTGATTTCTGGATCAGGGATACCACTTCTGTTCACACCCCGACCGAGGGCTAGAAA	1020		
Db	961	TGTTGATTTCTGGATCAGGGATACCACTTCTGTTCACACCCCGACCGAGGGCTAGAAA	1020		
Qy	1021	ATTGTTTGAGATTTTATATCATCTTGTCAAATGCTTCAGTTGTAATGTGAAAAATG	1080		

Db	1021	ATTGTTTGAGATTTTATATCATCTTGTCAAATGCTTCAGTTGTAATGTGAAAAATG	1080
Qy	1081	GGCTGGGAAAAGAGTGGTGCCTAAATGTTTTTACTTGTAACTTGTCTTGTGCCCC	1140
Db	1081	GGCTGGGAAAAGAGTGGTGCCTAAATGTTTTTACTTGTAACTTGTCTTGTGCCCC	1140
Qy	1141	TGGGCACCTTGGCCTTTGTCTCTCAGTGTCTTCCCTTTGACATGGGAAAGGATTGTG	1200
Db	1141	TGGGCACCTTGGCCTTTGTCTCTCAGTGTCTTCCCTTTGACATGGGAAAGGATTGTG	1200
Qy	1201	GCCAAAATCCCACTTCTTTGACACCTCAACGCTCTGTGGCTCAGGGCTGGGGTGGCAGAGG	1260
Db	1201	GCCAAAATCCCACTTCTTTGACACCTCAACGCTCTGTGGCTCAGGGCTGGGGTGGCAGAGG	1260
Qy	1261	GAGCCCTTCACTTATATCTGTGTTTATTCAGGGCTCCAGACTTCTCTCTGCGCTGC	1320
Db	1261	GAGCCCTTCACTTATATCTGTGTTTATTCAGGGCTCCAGACTTCTCTCTGCGCTGC	1320
Qy	1321	CCGACTGCACCTCTCCGCTTATCTATCTCTTCTCGGCTCCCGAGCCAGTCTTGGCT	1380
Db	1321	CCGACTGCACCTCTCCGCTTATCTATCTCTTCTCGGCTCCCGAGCCAGTCTTGGCT	1380
Qy	1381	TCTTGTCCGCTCTCGGGTCACTCCCTCCACTCTGACTCTGACTATGGCAGCAGAACCA	1440
Db	1381	TCTTGTCCGCTCTCGGGTCACTCCCTCCACTCTGACTCTGACTATGGCAGCAGAACCA	1440
Qy	1441	GGGCTGGCCAGTGGAATTCATGTTGATCAATTAAGAAAAATCGAACCAAAAAA	1500
Db	1441	GGGCTGGCCAGTGGAATTCATGTTGATCAATTAAGAAAAATCGAACCAAAAAA	1500
Qy	1501	AAAAA 1505	
Db	1501	AAAAA 1505	
RESULT 3			
AY358712	LOCUS	AY358712	1505 bp mRNA linear PRI 03-OCT-2003
DEFINITION	Homo sapiens clone DNA77626 ARP836 (UNQ836) mRNA, complete cds.		
ACCESSION	AY358712		
VERSION	AY358712.1 GI:37182544		
KEYWORDS	FLI CDNA.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 1505)		
AUTHORS	Clark,H.F., Gurney,A.L., Abaya,E., Baker,K., Baldwin,D., Brush,J., Chen,J., Chow,B., Chui,C., Crowley,C., Currell,B., Deuel,B., Dowd,P., Eaton,D., Foster,J., Grimaldi,C., Gu,Q., Haas,P.E., Heldens,S., Huang,A., Kim,H.S., Klimowski,L., Jin,Y., Johnson,S., Lee,J., Lewis,L., Liao,D., Mark,M., Robbie,E., Sanchez,C., Schoenfeld,J., Seshagiri,S., Simmons,L., Singh,J., Smith,V., Stinson,J., Vagts,A., Vandlen,R., Watanabe,C., Wieand,D., Woods,K., Xie,M.H., Yansura,D., Yi,S., Yu,G., Yuan,J., Zhang,M., Zhang,Z., Goddard,A., Wood,W.I. and Godowski,P.		
TITLE	The Secreted Protein Discovery Initiative (SPDI), a Large-Scale Effort to Identify Novel Human Secreted and Transmembrane Proteins: A Bioinformatics Assessment		
JOURNAL	Genome Res. 13 (10), 2265-2270 (2003)		
PUBMED	12975309		
REFERENCE	2 (bases 1 to 1505)		
AUTHORS	Clark,H.F.		
TITLE	Direct Submission		
JOURNAL	Submitted (01-AUG-2003) Department of Bioinformatics, Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080, USA		
FEATURES	Location/Qualifiers		
source	1..1505		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="DNA77626"		

```

gene      1. .1505
CDS       /locus_tag="UNQ836"
          112..894
          /locus_tag="UNQ836"
          /note="PRO1774"
          /codon_start=1
          /product="ARPG836"
          /protein_id="AAQ89074.1"
          /db_xref="GI:37182545"
          /translation="MARPMERWRDLALVTGASGGIGAAVARALVQGLKXVGCART
          VGNIEELAAECKSAGYPTGLPIYRCDLSNEEDILSMFSAIRSOHSGVDICINNAGLAR
          PDTLLSGSTGSKDMFNVLALISICTREAYSKMRNVDDGHIININSMGHRVLPL
          SVTHFYSAKVTALTEGLRELREAOHTIPATCISPGVVETQPAFLKHDKPEKAA
          ATYEQMKLCKPEDVAEYVYLSTPAHIQIGDIQMRPEQVT"
ORIGIN

```

```

Query Match      100.0%; Score 1505; DB 9; Length 1505;
Best Local Similarity 100.0%; Pred. No. 1.5e-278;
Matches 1505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGGATCGGACCCAAAGCAGGTCCGGCGGGCGGCGGAGAGAGCGCGGGCGTCAAGTCC 60
DB 1 CGCGGATCGGACCCAAAGCAGGTCCGGCGGGCGGCGGAGAGAGCGCGGGCGTCAAGTCC 60

QY 61 TCGACCCCGGTGTCGGGCTAGTCCAGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
DB 61 TCGACCCCGGTGTCGGGCTAGTCCAGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120

QY 121 CCCGCGATCGAGCGGTGCGCGGACCGGCTGCGGCTGGTGAAGGGGCGCTCGGGGGGCGATC 180
DB 121 CCCGCGATCGAGCGGTGCGCGGACCGGCTGCGGCTGGTGAAGGGGCGCTCGGGGGGCGATC 180

QY 181 GCGCGGCGGTGGCCCGGCGCTGGTCCAGCAGGCGACTGAAGGTGGTGGGCTGGCGCGCG 240
DB 181 GCGCGGCGGTGGCCCGGCGCTGGTCCAGCAGGCGACTGAAGGTGGTGGGCTGGCGCGCG 240

QY 241 ACTGTGGCAACATCGAGGAGTGGCTGCTGAATGAAGAGTGACGAGGTACCCCGGGACT 300
DB 241 ACTGTGGCAACATCGAGGAGTGGCTGCTGAATGAAGAGTGACGAGGTACCCCGGGACT 300

QY 301 TTGATCCCGCTACAGATGGAACCTATCAAAATGAAGAGGACATCCTCTCCATGTTCTCAGCT 360
DB 301 TTGATCCCGCTACAGATGGAACCTATCAAAATGAAGAGGACATCCTCTCCATGTTCTCAGCT 360

QY 361 ATCGCTTCTCAGCAGCGGTGTAGACATCTGCATCAACAATGCTGGCTTGGCGCGGCT 420
DB 361 ATCGCTTCTCAGCAGCGGTGTAGACATCTGCATCAACAATGCTGGCTTGGCGCGGCT 420

QY 421 GACACCCCTGCTCTCAGGCGAGCACCAGTGGTGGAGGACATGTTCAATGTGAACGTGCTG 480
DB 421 GACACCCCTGCTCTCAGGCGAGCACCAGTGGTGGAGGACATGTTCAATGTGAACGTGCTG 480

QY 481 GCCCTCAGCATCTGCAACCGGAAGCCTTACAGTCCATGAAGAGGCGGAAATGTGGACGAT 540
DB 481 GCCCTCAGCATCTGCAACCGGAAGCCTTACAGTCCATGAAGAGGCGGAAATGTGGACGAT 540

QY 541 GGGCAGCATCATTAACATCAATAGCATGTCGGCCACCGAGGTACCCCTCTCTGTGACC 600
DB 541 GGGCAGCATCATTAACATCAATAGCATGTCGGCCACCGAGGTACCCCTCTCTGTGACC 600

QY 601 CACTTCTATAGTGCACCAAGATATGCCGTCACTGCGCTGACAGAGGGGACTGAGGCAAGAG 660
DB 601 CACTTCTATAGTGCACCAAGATATGCCGTCACTGCGCTGACAGAGGGGACTGAGGCAAGAG 660

QY 661 CTTCGGAGGCGGACACCATCGAGCCACGTGTCATCTCTCCAGGTGGTGGAGACA 720
DB 661 CTTCGGAGGCGGACACCATCGAGCCACGTGTCATCTCTCCAGGTGGTGGAGACA 720

QY 721 CAATTCGCGCTTCAAACTCCAGCAAGGACCCCTGAGAAGGCGAGTGCACCTATGAGCAA 780
DB 721 CAATTCGCGCTTCAAACTCCAGCAAGGACCCCTGAGAAGGCGAGTGCACCTATGAGCAA 780

QY 781 ATGAAGTGTCTCAAAACCCGAGGATGTGGCGAGGCTGTTATCTACGTCCTCAGCACCCCC 840

```

```

Db 781 ATGAAGTGTCTCAAAACCCGAGGATGTGGCGAGGCTGTTATCTACGTCCTCAGCACCCCC 840
QY 841 GCACACATCCAGATTGGAGACATCCAGATGAGGCCCGCAGGAGCGGTGACCTAGTGACTG 900
Db 841 GCACACATCCAGATTGGAGACATCCAGATGAGGCCCGCAGGAGCGGTGACCTAGTGACTG 900
QY 901 TGGGAGCTCTCTCCCTCCCTCCCGACCTTCATGGCTTGGCTTCCCTCTCTGATTTTAGG 960
Db 901 TGGGAGCTCTCTCCCTCCCTCCCGACCTTCATGGCTTGGCTTCCCTCTCTGATTTTAGG 960
QY 961 TGTGATTCTTGGATCAGGGATACCACTTCTCTGTCACACCCCGCAGGCGGCTAGAAA 1020
Db 961 TGTGATTCTTGGATCAGGGATACCACTTCTCTGTCACACCCCGCAGGCGGCTAGAAA 1020
QY 1021 ATTTGTTTGAGATTTTATATCATCTTGTCAAAATGCTTCAGTTGTAATGTGAAAAATG 1080
Db 1021 ATTTGTTTGAGATTTTATATCATCTTGTCAAAATGCTTCAGTTGTAATGTGAAAAATG 1080
QY 1081 GGCTGGGAAAAAGAGGTGGTCCCTAAATGTTTACTTGTAACTTGTCTTGTGCCCC 1140
Db 1081 GGCTGGGAAAAAGAGGTGGTCCCTAAATGTTTACTTGTAACTTGTCTTGTGCCCC 1140
QY 1141 TGGGCACATTGGCTTTGCTCTCTCAGTGTCTTCCCTTTGACATGGGAAAGAGTTGTG 1200
Db 1141 TGGGCACATTGGCTTTGCTCTCTCAGTGTCTTCCCTTTGACATGGGAAAGAGTTGTG 1200
QY 1201 GCCAAAATCCCATCTTCTTGCACCTCAACCTGTGTGGCTCAGGCTCGGGTGGCAGAGG 1260
Db 1201 GCCAAAATCCCATCTTCTTGCACCTCAACCTGTGTGGCTCAGGCTCGGGTGGCAGAGG 1260
QY 1261 GAGGCTTCACCTTATATCTGTGTGTTATCCAGGGCTCAGACTTCTCTCTGCTGCTGC 1320
Db 1261 GAGGCTTCACCTTATATCTGTGTGTTATCCAGGGCTCAGACTTCTCTCTGCTGCTGC 1320
QY 1321 CCCACTGCACCTCTCCCTTATCTATCTCTTCTCGGCTCCCGAGCCAGCTCTTGCT 1380
Db 1321 CCCACTGCACCTCTCCCTTATCTATCTCTTCTCGGCTCCCGAGCCAGCTCTTGCT 1380
QY 1381 TCTTGTCCCTCTCGGGTCACTCCCTCACTCTGACTCTGACTATGGCAGCAGAACACCA 1440
Db 1381 TCTTGTCCCTCTCGGGTCACTCCCTCACTCTGACTCTGACTATGGCAGCAGAACACCA 1440
QY 1441 GGGCTCGCCCGAGTGGATTTTCATGCTATCAATTAAGAAAGAAATCCACACCAAAAA 1500
Db 1441 GGGCTCGCCCGAGTGGATTTTCATGCTATCAATTAAGAAAGAAATCCACACCAAAAA 1500
QY 1501 AAAAA 1505
Db 1501 AAAAA 1505

RESULT 4
BC002731
LOCUS      1501 bp      mRNA      linear      PRI 12-NOV-2003
DEFINITION Homo sapiens hypothetical protein MGC4172, mRNA (cdna clone
MGC:4172 IMAGE:3631177), complete cds.
ACCESSION  BC002731
VERSION     BC002731.2  GI:33877092
KEYWORDS   MGC.
SOURCE      Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 1501)
            Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
            Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
            Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
            Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
            Datchenko, L., Narusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
            Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
            Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
            Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,

```

Abramson, R.D., Mullaby, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Butterfield, Y.S., Krywinski, M.I., Skalska, U., Smalrus, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A., 2003. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. *Proc. Natl. Acad. Sci. U.S.A.* 99 (26), 16899-16903 (2002) 22388257

2 (bases 1 to 1501)

Strausberg, R.

Direct Submission

Submitted (05-FEB-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

On Aug 19, 2003 this sequence version replaced gi:12803782.

Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: nisc.mgc@nih.gov

Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.I., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stauripop, S., Thomas, P.J., Touchman, J.W., Tsurgone, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>

Series: IRAL Plate: 12 Row: g Column: 18

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 13236541.

location/Qualifiers

1. .1501

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:4172 IMAGE:3631177"

/tissue_type="uterus, endometrium adenocarcinoma"

/clone_lib="NIH MGC 44"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

1. .1501

/gene="MGC4172"

/note="synonym: FLJ22543"

/db_xref="LocusID:79154"

338. .883

/codon_start=1

/product="hypothetical protein MGC4172"

/protein_id="AAH02731.1"

/db_xref="GI:12803783"

/db_xref="LocusID:79154"

/translation="MPSAIPRSHQVDICINNAAGLAPDPFLLSGTSQWKMENVNLALSICTREAYQSMKERNVDGHLININSMGHRVFLSVTHFYSAKYAVTALTEGLRQELREAGTHIRATCISPGVVEVTFQAFKLHKDPEKAAATYEQMKCLKPEDVAEVIYVLSPTAHIOIGDIOQMRPEQVT"

338. .832

misc_feature

/note="adh_short; Region: short chain dehydrogenase. This family contains a wide variety of dehydrogenases"

/db_xref="CDD:pfam00106"

Query Match 99.3%; Score 1494; DB 9; Length 1501; Best Local Similarity 100.0%; Pred. No. 1.9e-276; Matches 1494; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN

QY 12 CCCAAGCAGGTGCGGGCGGCGGAGGAGCGGCGGCGGTCTCAGCTCTCGACCCCGCT 71

Db 1 CCCAAGCAGGTGCGGGCGGCGGAGGAGCGGCGGCGGTCTCAGCTCTCGACCCCGCT 60

QY 72 GTCGGGCTAGTCCAGCAGGCGGAGCGGCGGCGGTCTCAGCTCTCGACCCCGCT 131

Db 61 GTCGGGCTAGTCCAGCAGGCGGAGCGGCGGCGGTCTCAGCTCTCGACCCCGCT 120

QY 132 GCGGTGGCGGCGGCGGCGGAGGAGCGGCGGCGGTCTCAGCTCTCGACCCCGCT 191

Db 121 GCGGTGGCGGCGGCGGCGGAGGAGCGGCGGCGGTCTCAGCTCTCGACCCCGCT 180

QY 192 GCGCGGCGGCGGCGGCGGAGGAGCGGCGGCGGTCTCAGCTCTCGACCCCGCT 251

Db 181 GCGCGGCGGCGGCGGCGGAGGAGCGGCGGCGGTCTCAGCTCTCGACCCCGCT 240

QY 252 CATCGAGGAGTGGCTGCTGAATGTGAAGTGCAGGCTACCCCGGAGCTTTGATCCCTCA 311

Db 241 CATCGAGGAGTGGCTGCTGAATGTGAAGTGCAGGCTACCCCGGAGCTTTGATCCCTCA 300

QY 312 CAGATGTACCTATCAATGAAGGAGCATCTCTCCATGTTCTCAGTATCCGTTCTCA 371

Db 301 CAGATGTACCTATCAATGAAGGAGCATCTCTCCATGTTCTCAGTATCCGTTCTCA 360

QY 372 GCACAGCGGTGTAGACATCTGCATCAACAAGTGGCTTGGCGGCGGTGACACCTGCT 431

Db 361 GCACAGCGGTGTAGACATCTGCATCAACAAGTGGCTTGGCGGCGGTGACACCTGCT 420

QY 432 CTCAGGAGCAGCAGTGGTTCGAAGGAGCATGTTCAATGTGAACGTGCTGGCGCTCAGCAT 491

Db 421 CTCAGGAGCAGCAGTGGTTCGAAGGAGCATGTTCAATGTGAACGTGCTGGCGCTCAGCAT 480

QY 492 CTGCACACGGGAAGCCTACCAATCCATGAAGGAGGAGTGTGACGATGGGACATCAT 551

Db 481 CTGCACACGGGAAGCCTACCAATCCATGAAGGAGGAGTGTGACGATGGGACATCAT 540

QY 552 TAAATCAATAGCATGTCTGGCCACCGAGTGTATACCTGCTGTGACCCATCTCTATAG 611

Db 541 TAAATCAATAGCATGTCTGGCCACCGAGTGTATACCTGCTGTGACCCATCTCTATAG 600

QY 612 TGCACCAAGTATGCCCTCTACCTGCGCTGACAGAGGAGCTGAGGCAAGAGCTTCGGGAGGC 671

Db 601 TGCACCAAGTATGCCCTCTACCTGCGCTGACAGAGGAGCTGAGGCAAGAGCTTCGGGAGGC 660

QY 672 CCAGACCCACATCCGAGCCACGTCATCTCTCAGGTGTGGTGGAGACACAATTCGCGCTT 731

Db 661 CCAGACCCACATCCGAGCCACGTCATCTCTCAGGTGTGGTGGAGACACAATTCGCGCTT 720

QY 732 CAAACTCCACGACAAAGGAGCCTTGAGAGGAGCTGCCACCTATGACCAATGAAGTGTCT 791

Db 721 CAAACTCCACGACAAAGGAGCCTTGAGAGGAGCTGCCACCTATGAGCAATGAAGTGTCT 780

QY 792 CAAACCGGAGATGTGGCGGAGCTGTTATCTACGTCCTCAGCACCCCGCCACACATCCA 851

Db 781 CAAACCGGAGATGTGGCGGAGCTGTTATCTACGTCCTCAGCACCCCGCCACACATCCA 840

QY 852 GATTGGAGACATCCAGATGAGGCCACGAGGAGGTGACCTAGTGAATGTGGAGTCTCT 911

Db 841 GATTGGAGACATCCAGATGAGGCCACGAGGAGGTGACCTAGTGAATGTGGAGTCTCT 900

QY 912 CTTTCCCTCCCAACCTTCTATGCTTGCCTCTCTGCTCTGGAATTTAGTGTGATTTCT 971

Db 901 CTTTCCCTCCCAACCTTCTATGCTTGCCTCTCTGCTCTGGAATTTAGTGTGATTTCT 960

Qy	972	GGATCAGGGGATACCACTTCTCTGTCCACACCCCGACAGGGGCTAGAAAATTTGTTGAG	1031
Db	961	GGATCAGGGGATACCACTTCTCTGTCCACACCCCGACAGGGGCTAGAAAATTTGTTGAG	1020
Qy	1032	ATTTTATATCATCTTGTCAAAATGCTTCACTGTAAATGTGAAAATGGGCTGGGAAA	1091
Db	1021	ATTTTATATCATCTTGTCAAAATGCTTCACTGTAAATGTGAAAATGGGCTGGGAAA	1080
Qy	1092	GGAGGTGGTGCCTCAATATGTTTACTTGTAACTTGTGTGTCGCCCTGGGCACTGG	1151
Db	1081	GGAGGTGGTGCCTCAATATGTTTACTTGTAACTTGTGTGTCGCCCTGGGCACTGG	1140
Qy	1152	CCTTTGCTGTCTCAGTGTCTTCCCTTTGACATGGGAAGAGTGTGGCCAAATCCC	1211
Db	1141	CCTTTGCTGTCTCAGTGTCTTCCCTTTGACATGGGAAGAGTGTGGCCAAATCCC	1200
Qy	1212	CATCTTCTTGCACTCAACGCTCTGTGGCTCAGGGCTGGGGTGGCAGAGGAGCCCTTCAC	1271
Db	1201	CATCTTCTTGCACTCAACGCTCTGTGGCTCAGGGCTGGGGTGGCAGAGGAGCCCTTCAC	1260
Qy	1272	CTTATATCTGTGTGTTATCCAGGCTCCAGACTTCTCTCTGCTGCTGCCCTGACACC	1331
Db	1261	CTTATATCTGTGTGTTATCCAGGCTCCAGACTTCTCTCTGCTGCTGCCCTGACACC	1320
Qy	1332	CTCTCCCTTATCTATCTCTCTTCTGGCTCCCCAGCCAGTCTTGGCTTCTTGTCCCTT	1391
Db	1321	CTCTCCCTTATCTATCTCTTCTGGCTCCCCAGCCAGTCTTGGCTTCTTGTCCCTT	1380
Qy	1392	CCTGGGCTCATCCCTCCACTCTGACTCTGACTATGGCAGCAGAACACAGGGCTGGCCC	1451
Db	1381	CCTGGGCTCATCCCTCCACTCTGACTCTGACTATGGCAGCAGAACACAGGGCTGGCCC	1440
Qy	1452	AGTGGATTTTCATGGTGATCATTAATAAGAAAAATCGACCAAAAAA	1505
Db	1441	AGTGGATTTTCATGGTGATCATTAATAAGAAAAATCGACCAAAAAA	1494
RESULT 5			
AX179289			
LOCUS	AX179289	1909 bp	DNA linear PAT 03-JUL-2001
DEFINITION	Sequence 2 from Patent WO0144446.		
ACCESSION	AX179289		
VERSION	AX179289.1 GI:14598959		
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1		
AUTHORS	Meyers, P.		
TITLE	21612, 21615, 21620, 21676, 33756, novel human alcohol dehydrogenases		
JOURNAL	Patent: WO 0144446-A 2 21-JUN-2001; Millennium Pharmaceuticals, Inc. (US)		
FEATURES	Location/Qualifiers		
	1..1909		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
CDS	/db_xref="taxon:9606"		
	421..1203		
	/note="unnamed protein product; 21620 ADH"		
	/codon_start=1		
ORIGIN	/protein_id="CAC43880.1"		
	/db_xref="GI:14598960"		
	/translation="MARPGMRWRDLALVTGSGGIGAAVARALVQOGLKVVGCART VGNTEELAAECKSAGYPTLI PYRCLDSNEEDILSMFSAIRSOHSVDICINNAGLAR PDLTLLSGTSKGMKNVNLALSICTREAYOSKERNVDDGHIININMSGHRLVPL SVTHFYSATKVAIVTALTGELGRQELREAQTHIRATCISPGVVEVTFQFAKLHDKDPEKAA ATYEQMKCLPEDVAEAVIYVLSTPAHIQIGIDIQMRPEQVT"		
Query Match	99.1%	Score 1492; DB 6; Length 1909;	

Best Local Similarity 99.8%; Pred. No. 4.5e-276;			
Matches 1503; Conservative 1; Mismatches 1; Indels 1; Gaps 1;			
Qy	1	CGCGGATCGGACCCAAAGCAGGTCCGCGCGCGCGCAGGAGCGCGCGGCGTCCAGTCC	60
Db	310	CGCGGATCGGACCCAAAGCAGGTCCGCGCGCGCGCAGGAGCGCGCGGCGTCCAGTCC	369
Qy	61	TCGACCCCGGTGTCGGGCTAGTCCAGCGAGCGCGCGCGCGTGGGCCCATGGCCAGG	120
Db	370	TCGACCCCGGTGTCGGGCTAGTCCAGCGAGCGCGCGCGCGTGGGCCCATGGCCAGG	429
Qy	121	CCCGGATGGAGCGGTGCGCGACCGGTGCGCTGTGTGACGGGGGCTCCGGGGGGCATC	180
Db	430	CCCGGATGGAGCGGTGCGCGACCGGTGCGCTGTGTGACGGGGGCTCCGGGGGGCATC	489
Qy	181	GGCGGGCGGTGGCCCGGGCCCTGGTCAGCAGGAGCTGAAAGGTGGTGGGCTGGCCCGC	240
Db	490	GGCGGGCGGTGGCCCGGGCCCTGGTCAGCAGGAGCTGAAAGGTGGTGGGCTGGCCCGC	549
Qy	241	ACTGTGGGCAACATCGAGGAGCTGGCTGCTCAATGTAAAGTGCAGGCTACCCCGGGACT	300
Db	550	ACTGTGGGCAACATCGAGGAGCTGGCTGCTCAATGTAAAGTGCAGGCTACCCCGGGACT	609
Qy	301	TTGATCCCGCTACAGATGTGACCTATCAAAATGAAGAGGACATCCTCTCCATGTTCTCAGCT	360
Db	610	TTGATCCCGCTACAGATGTGACCTATCAAAATGAAGAGGACATCCTCTCCATGTTCTCAGCT	669
Qy	361	ATCGGTTCTCAGCAGCGGTGTAGACATCTGCATCAACATGCTGGCTTGGCCCGGCT	420
Db	670	ATCGGTTCTCAGCAGCGGTGTAGACATCTGCATCAACATGCTGGCTTGGCCCGGCT	729
Qy	421	GACACCTGTCTCTCAGGCGAGCACCACTGCTTGGAGGACATGTTCAATGTGAACGTGTG	480
Db	730	GACACCTGTCTCTCAGGCGAGCACCACTGCTTGGAGGACATGTTCAATGTGAACGTGTG	789
Qy	481	GCCCTCAGCATCTGCACACGGGAAGCCTACCACTGCTCAATGAAGAGCGGAATGTGACGAT	540
Db	790	GCCCTCAGCATCTGCACACGGGAAGCCTACCACTGCTCAATGAAGAGCGGAATGTGACGAT	849
Qy	541	GGGACATCATTAACATCAATAGCATGTCTGSCCACCAGTGTATACCCCTGTCTGTGACC	600
Db	850	GGGACATCATTAACATCAATAGCATGTCTGSCCACCAGTGTATACCCCTGTCTGTGACC	909
Qy	601	CACCTTCTATAGTGCCACCAAGTATGCCGTCACTGCGCTGACAGAGGACTGAGGCAAGAG	660
Db	910	CACCTTCTATAGTGCCACCAAGTATGCCGTCACTGCGCTGACAGAGGACTGAGGCAAGAG	969
Qy	661	CTTCGGGAGGCCAGACCCACATCCGAGCCACGTGCACTCTCCAGGTGTGTGGGAGACA	720
Db	970	CTTCGGGAGGCCAGACCCACATCCGAGCCACGTGCACTCTCCAGGTGTGTGGGAGACA	1029
Qy	721	CAATTGCGCTTCAAACTCCACGACAGGACCTTGAGAGGAGCTGCCACCTATGAGCAA	780
Db	1030	CAATTGCGCTTCAAACTCCACGACAGGACCTTGAGAGGAGCTGCCACCTATGAGCAA	1089
Qy	781	ATGAAGTGTCTCAAAACCCGAGGATGTGGCCGAGGCTGTTATCTACGCTCTCAGACCCCC	840
Db	1090	ATGAAGTGTCTCAAAACCCGAGGATGTGGCCGAGGCTGTTATCTACGCTCTCAGACCCCC	1149
Qy	841	GCACACATCCAGATTGGAGACATCCAGATGAGGCCACCGAGGAGTGTAGTACTG	900
Db	1150	GCACACATCCAGATTGGAGACATCCAGATGAGGCCACCGAGGAGTGTAGTACTG	1209
Qy	901	TGGGAGCTCTCTCTCCCTCCACCTTTCATGGCTTGCCTCTCTGCTCTCGATTTTAGG	960
Db	1210	TGGGAGCTCTCTCTCTCCCTCCACCTTTCATGGCTTGCCTCTCTGCTCTCGATTTTAGG	1269
Qy	961	TGTTGATTTCTGGATCAGCGGATACCACTTCTCTGTCCACACCCCGACCGAGGCGCTAGAAA	1020
Db	1270	TGTTGATTTCTGGATCAGCGGATACCACTTCTCTGTCCACACCCCGACCGAGGCGCTAGAAA	1329
Qy	1021	ATTGTTTGGATTTTATATCATCTGTGCAAAATGCTTTCAGTTGTAATGTGAAAATG	1080

Db	837	CTCCCTCTGGATTTTAGTGTGATTTCTGGATACCGGGATACCACTTCTCTGCCACACC	896
Qy	1003	CCGACACAGGGCTAGAAAAATTTTTCAGATTTTATATCATCTTGTCAAATGCTTTCAG	1062
Db	897	CCGACACAGGGCTAGAAAAATTTTTCAGATTTTATATCATCTTGTCAAATGCTTTCAG	956
Qy	1063	TTGTAAATGTGAAAAATGGGCTGGGAAAGAGAGTGTGTCCCTAAATGTTTACTTGT	1122
Db	957	TTGTAAATGTGAAAAATGGGCTGGGAAAGAGAGTGTGTCCCTAAATGTTTACTTGT	1016
Qy	1123	AACCTGTCTGTGCTCCCTGGGCACTTGGCCCTTGTCTGCTCTCAGTGTCTTCCCTTTGA	1182
Db	1017	AACCTGTCTGTGCTCCCTGGGCACTTGGCCCTTGTCTGCTCTCAGTGTCTTCCCTTTGA	1076
Qy	1183	CATGGGAAAGAGTGTGGCCAAATCCCATCTTCTTGGCACTCAACGCTGTGGCTCA	1242
Db	1077	CATGGGAAAGAGTGTGGCCAAATCCCATCTTCTTGGCACTCAACGCTGTGGCTCA	1136
Qy	1243	GGGCTGGGGTGGAGAGGAGGCTTACCTTATATCTGTGTTTGTATCCAGGCTCCAG	1302
Db	1137	GGGCTGGGGTGGAGAGGAGGCTTACCTTATATCTGTGTTTGTATCCAGGCTCCAG	1196
Qy	1303	ACTTCTCTCTGCTGCTCCCTGACCTGACCTCTCCCTTATCTATCTATCTCTCTCGGCTC	1362
Db	1197	ACTTCTCTCTGCTGCTCCCTGACCTGACCTCTCCCTTATCTATCTATCTCTCGGCTC	1256
Qy	1363	CCGACCCAGCTTGGCTTCTTGTCCCTCTCTGGGCTCATCCCTCCATCTGACTCTGAC	1422
Db	1257	CCGACCCAGCTTGGCTTCTTGTCCCTCTCTGGGCTCATCCCTCCATCTGACTCTGAC	1316
Qy	1423	TATGGCAGCAGAACACAGGGGCTGGCCAGTGGATTTTCATGGTGCATTTAAAAAGAA	1482
Db	1317	TATGGCAGCAGAACACAGGGGCTGGCCAGTGGATTTTCATGGTGCATTTAAAAAGAA	1376
Qy	1483	AAATCGCACCAAAAAA	1505
Db	1377	AAATCGCACCAAAAAA	1399
RESULT 7			
LOCUS	BC022224		
DEFINITION	Mus musculus cDNA sequence BC022224, mRNA (cdna clone MGC:18716 IMAGE:4219994), complete cds.		
ACCESSION	BC022224		
VERSION	BC022224.1 GI:18490589		
KEYWORDS	MGC.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE			
AUTHORS	Klausner, R.D., Collins, F.S., Wagner, L., Shemen, C.M., Schuler, G.D., Altshul, S.F., Zebner, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Donald, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, R.J., Malek, J.A., Gunaratne, P.H., Richards, S., Wille, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallus, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.		
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)		

MEDLINE	22388257		
PUBMED	12477932		
REFERENCE	2 (bases 1 to 1398)		
AUTHORS	Strausberg, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (01-FEB-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov		
COMMENT	Contact: MGC help desk Email: cgapsb@mail.nih.gov Tissue Procurement: Jeffrey E. Green, M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL) DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: http://www-shgc.stanford.edu Contact: (Dickson, Mark) mcd@paxil.stanford.edu Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.		
FEATURES	source		
	1..1398		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/strain="FVB/N"		
	/db_xref="taxon:10090"		
	/clone="MGC:18716 IMAGE:4219994"		
	/tissue_type="Colon, normal, 5 month old male mouse."		
	/clone_lib="NCI CGAP_C024"		
	/lab_host="DH10B"		
	/note="Vector: pCMV-SPORT6"		
gene	1..1398		
	/gene="BC022224"		
	/note="synonym: MGC18716"		
	/db_xref="LocusID:192970"		
	/db_xref="MGI:2652816"		
CDS	305..850		
	/codon_start=1		
	/product="cDNA sequence BC022224"		
	/protein_id="AAH22224.1"		
	/db_xref="GI:19490590"		
	/db_xref="LocusID:192970"		
	/translation="MFSVRSQHSQSGVDCINNAGMARPDLLSGTSGMKMFNNVNL ALSICTRAYOSMKERNIDDDGHIINSMCGHRVPPQSVIHFYSATKYATVATLEGLR QELLEAQTHIRATCISPLGVETQFAFLHDKDPGEAAATYEHKLRPEDVAEAVIY LSTPPHVQVQVDIQMRPTQVT"		
misc_feature	305..844		
	/note="COG4221; Region: Short-chain alcohol dehydrogenase of unknown specificity [General function prediction only]"		
	/db_xref="CDD:COG4221"		
ORIGIN			
	Query Match 50.3%; Score 756.4; DB 10; Length 1398;		
	Best Local Similarity 76.2%; Pred. No. 6.3e-135;		
	Matches 1007; Conservative 0; Mismatches 286; Indels 29; Gaps 5;		
Qy	46	CCGGGGGTGAGTCTCTCGACCCCGTGTGGGTAGTCCAGCGAGCGGCGCGCGT	105
Db	2	CCGAGCTTAAGTCCCGGTCCAGGCTTCCGAAACAGAGTGGCAGACGCGTC	61
Qy	106	GGSCCCATGCCAGGCGCGGCGATGGAGCGGTGGCGGACCGGCTGGTGGACGGG	165
Db	62	GGGACCATGACTAGAGCTGGCATGGAGCGGTGGCGGACCGGCTGGCATGGT	121
Qy	166	GCCTCGGGGGGCATCGGCGCGCGCTGGCCCGGCGCTGTCCAGCAGGACTGA	225

Db 122 GCTCGGGGGGATCGGTGGCGCGGTGGCCGGGCATTAGTCCAGCAGGACTGAAGGT 181
QY 226 GTGGGCTGGCGCGGACATGTTGGGCGAATCGAGGAGCTGGCTGTAATGTAAGAGTGCA 285
Db 182 GTGGGTTGTGGCGCGGACCGTTGGCAACATCGAGGAGCTGGCTGCTGAATGTAAGAGTGCA 241
QY 286 GGCTACCCCGGAGCTTTGATCCCTTACAGATGTGACCTTATCAAAATGAAGAGGACATCCTC 345
Db 242 GGCTACCCCGGAGCTTTGATCCCTTACAGATGTGACCTTATCAAAATGAAGAGGACATCCTC 301
QY 346 TCCATGTTCTAGCTATCGTTTCTCAGCACACGGGTAGACATCTGCATCAACAATGCT 405
Db 302 TCCATGTTCTAGCTATCGTTTCTCAGCACACGGGTAGACATCTGCATCAACAATGCT 361
QY 406 GCTTGGCGCGGCTCGACACCTGCTCTCAGCAGCAGCAGCTGGTGTGAAGAGACATGTT 465
Db 362 GGCATGGCGCGGCTCGACACCTGCTCTCAGCAGCAGCAGCTGGTGTGAAGAGACATGTT 421
QY 466 AATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCTACAGTCCATGAAGAG 525
Db 422 AATGTGAATGTGCTGGCCCTCAGCATCTGCATCGGAGGCTTATCAGTCCATGAAGAG 481
QY 526 CGGATGTGGAGGATGGGCAATCATTAATCAATGATCTCTGGCCACCGAGTGTTA 585
Db 482 CGGAACATAGACGAGGCGCATCATTAATCAATGATCTCTGGCCACCGAGTGTTA 541
QY 586 CCCCTGTCTGTGACCACTCTATAGTGCACCAAGTATGCGTCACTGCCTGACAGAG 645
Db 542 CCCAGTCTGTGATCCATTTCTATAGTGCAGTAAAGTATGCGTCACTGCCTGACAGAG 601
QY 646 GGACTGAGGCAAGAGCTTTCGGGAGGCGCCAGACCCACATCCAGGACCATGCTCTTCCA 705
Db 602 GGACTGAGGCAAGAGCTTTCGGGAGGCGCCAGACCCATATCCGGGCCCATGTTCTTCCA 661
QY 706 GGTGTGTGGAGACATATGCTTCTTAACTCCAGCAAGAGACCTGTGAAGGAGAGCT 765
Db 662 GGTGTGTGGAGACATATGCTTCTTAACTCCAGCAAGAGACCTGTGAAGGAGAGCT 721
QY 766 GCCACCTATGAGCAATGAAGTGTCTCAACCCAGGATGTCGGCGAGGCTGTTATCTAC 825
Db 722 GCCACCTATGAGCAATGAAGTGTCTCAACCCAGGAGCTGTCGAGGCTGTTATCTAC 781
QY 826 GTCTCAGACCCCGGACACATCCAGATGGAGACATCCAGATGAGGCGCCACGAGGAG 885
Db 782 GTTCTTAGCACACCCACATGTTTCAAGTGTGTTGATCCAGATGAGGCGCCACGAGGAG 841
QY 886 GTGACCTAGTGTGGGAGCTCTCTTCCCTCCGACCCCTTCAATGCTTGCCTCTG 945
Db 842 GTGACCTAGTGTGGGAGCTCTCTTCCCTCCGACCCCTTCAATGCTTGCCTCTG 900
QY 946 CCTCTGGAATTTAGGTTGATTTCTGATCAGGATCAGGATACCATCTCTGTCACACCCG 1005
Db 901 CCTCTGGAATTTAGGTTGATTTCTGGA-CTCTGAGCTTACTTTCTATGTCCACCCCA 959
QY 1006 ACCAGGGGTAGAAAAATTTGTTGAGATTTTATATCATCTGTGCAATGCTTCAGTTG 1065
Db 960 CTAAGGG-----ATAGACAGCTTGTACAGCTATTCCCATTA 995
QY 1066 TAATGTGAATAATGGCTGGGAAGAGGAGTGGTGTCCCTAATGTTTACTGTGTTAAC 1125
Db 996 CAAATGTGAATAATGGCTGGGAAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1055
QY 1126 TTGTTCTTGTGCGCCCTGGGACATCTGGCTTGTGCTCTCAGTGTCTTCCCTTTGACAT 1185
Db 1056 TTCTGAGTCTCTTGGGCTCCGTAACTTTG-CTCTCTGTGCGCTTCTCTTTGACCT 1114
QY 1186 GGAAGAGGAGTGTGGGCAAAATCCCATCTTCTTGCACTCAACCTCTGTGTGGCTCAGG 1245
Db 1115 AGGAAGAGGAGTGTACCTAGAGGGGAGTGTCTCTATCTCTTACACCTTACCCCTCAGCC 1174
QY 1246 CT--GGGTGGGAGAGGAGGCTTCACTTATATCTGTGTTGTTATCCAGGCTCCAGA 1303

Db 1175 CTTAGGGTTGGAGCAGCAGACCTCATCTTTTACCTGTGTGTTTCCAGGGTTCCAGG 1234
QY 1304 CTTCTCTCTCTGCTGCCCCCACTGACACCTCTTCCCTTATCTATCTCTCTTCCGCTCC 1363
Db 1235 TTCCCTTCCCAACCTCTGCTCTCCCCCCCCCAGCTCCAGCAAGTTTGGCCCTC 1294
QY 1364 CC 1365
Db 1295 AC 1296

RESULT 8
AC003042/c
LOCUS AC003042 102818 bp DNA linear PRI 14-JUL-1998
DEFINITION Homo sapiens chromosome 17, clone HCIT75G16, complete sequence.
AC003042
AC003042.1 GI:3319121
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 102818)
AUTHORS Birren, B., Fasman, K., Linton, L., Nusbaum, C. and Lander, E.
JOURNAL Homo sapiens chromosome 17, clone HCIT75G16
Unpublished
REFERENCE 2 (bases 1 to 102818)
AUTHORS Birren, B., Fasman, K., McKernan, K., Munro, C., Nusbaum, C.,
Richardson, P., Lander, E., Baldwin, J., Barna, N., Cantu, C., Chang, A.,
Cooke, P., Daly, M. J., Devon, K., Dewar, K., Duret, B., Forrest, C.,
Gage, D., Gensheimer, S., Geraigery, K., Gilmartin, T., Hags, B.,
Halphen, J., Harris, K., Howland, J. C., Huang, J., Hui, L., Jacotot, L.,
Karby, A., Lane, M., Mackenzie, J., Marquis, N., McDermott, J.,
Molla, M., Morrow, J., Nachman, A., Naylor, J., O'Connor, T., Olotu, A.,
Peterson, K., Roberts, D., Rollins, G., Sarnaik, A., Shiu, P., Shyam, R.,
Stillwell, J., Stone, C., Strickland, C., Sydney, K., Tang, L.,
Zentseva, I. and Zody, M.
Direct Submission
Submitted (31-OCT-1997) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to 102818)
AUTHORS Birren, B., Fasman, K., Linton, L., Nusbaum, C., Lander, E., Allen, N.,
Baker, J., Baldwin, J., Barna, N., Beckerly, R., Benn, J., Boatman, C.,
Boutwell, C., Brown, A., Castle, A., Cerny, J., Cooke, P., Depayre, E.,
Devon, K., Dewar, K., Donegan, L., Etemadi, S., Ferreira, P.,
FitzHugh, W., Forrest, C., Funke, R., Gage, D., Gardyna, S.,
Gensheimer, S., Geraigery, K., Gilmartin, T., Grant, G., Hags, B.,
Harris, K., Horton, L., Howland, J. C., Hui, L., Jacotot, L., Kann, L.,
Macdonald, P., Marquis, N., Morris, W., Morrow, J., Mychaleckyj, J.,
Meldrum, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J.,
Nachman, A., Naylor, J., Naylor, J., Niloff, M., O'Connor, T., Pavlin, B.,
Peterson, K., Riley, R., Roberts, D., Rossello, R., Roy, A., Shyam, R.,
Stange-Thomann, N., Stillwell, J., Stojanovic, N., Stone, C., Strickland, C.,
Strickland, C., Subramanian, A., Torruella-Miller, I., Vassiliev, H.,
Vo, A., Wagner, A., Wang, B., Wheeler, J., Wu, Y., Ye, W. J., Zhao, J. and
Zody, M.
Direct Submission
Submitted (14-JUL-1998) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Jul 14, 1998 this sequence version replaced gi:3294535.
All repeats were identified using RepeatMasker: Smit, A.F.A. &
Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html.
FEATURES
Location/Qualifiers
1. .102818
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="HCIT75G16"
/clone_lib="Research Genetics/Cal Tech CITB978SK-B (plates
1-194)"

Fri Sep 17 09:20:43 2004

```
repeat_region 495. .778
/rpt_family="AluJb"
complement(1357. .1409)
/rpt_family="AT_rich"
repeat_region 1737. .1800
/rpt_family="CA)n"
repeat_region 1911. .1964
/rpt_family="MIR"
repeat_region 2074. .2216
/rpt_family="MIR"
repeat_region 2268. .22558
/rpt_family="AluSx"
complement(22631. .22688)
/rpt_family="TIGER2"
repeat_region 23162. .23460
/rpt_family="AluSx"
repeat_region 23461. .23512
/rpt_family="GGA)n"
repeat_region 23515. .23637
/rpt_family="GGA)n"
repeat_region 23761. .23761
/rpt_family="L2"
repeat_region 26652. .26652
/rpt_family="AluSx"
repeat_region 26941. .27047
/rpt_family="CAGA)n"
repeat_region 28673. .29257
/rpt_family="L1MB4"
complement(29262. .29556)
/rpt_family="AluSx"
repeat_region 29562. .30018
/rpt_family="L1MB4"
repeat_region 30035. .30160
/rpt_family="MIR"
repeat_region 30320. .30669
/rpt_family="L2"
repeat_region 31247. .31550
/rpt_family="AluSx"
complement(31605. .31741)
/rpt_family="MIR"
repeat_region 31919. .32211
/rpt_family="AluJb"
complement(32934. .33227)
/rpt_family="AluSx"
complement(33268. .33560)
/rpt_family="AluSx"
complement(34210. .34319)
/rpt_family="MIR"
repeat_region 34835. .34869
/rpt_family="CAT)n"
complement(34876. .35054)
/rpt_family="MIR"
complement(36108. .36327)
/rpt_family="MIR"
complement(36957. .37440)
/rpt_family="MLT2CB"
repeat_region 37827. .38129
/rpt_family="AluSx"
complement(38139. .38249)
/rpt_family="MER5B"
repeat_region 38410. .38708
/rpt_family="AluY"
repeat_region 38730. .39009
/rpt_family="AluSx"
complement(39165. .39287)
/rpt_family="MIR"

Query Match 42.9%; Score 646; DB 9; Length 102818;
Best Local Similarity 100.0%; Pred. No. 8.2e-114;
Matches 646; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

850 CAGATTGGAGACATCCAGATGAGGCCCGGAGGACCTAGTGTGGAGCTC 909
```

```
Db 48792 CAGATTGGAGACATCCAGATGAGGCCCGGAGAGAGGTGACCTAGTACGTGTGGAGCTC 48733
Qy 910 CTCCTTCCCTCCCAACCCTTCATGGCTTGCTCTCTGCTCTCGAATTTAGGTGTGATTT 969
Db 48732 CTCCTTCCCTCCCAACCCTTCATGGCTTGCTCTCTGCTCTCGAATTTAGGTGTGATTT 48673
Qy 970 CTGGATCAGGGATACCACTTCCCTGTCACACCCGACCGGCTAGAAAAATTTGTTTG 1029
Db 48672 CTGGATCAGGGATACCACTTCCCTGTCACACCCGACCGGCTAGAAAAATTTGTTTG 48613
Qy 1030 AGATTTTATATCATCTTGTCAAAATTCCTTCAGTTGTAAATGTCAAAAATGGCGCTGGGGA 1089
Db 48612 AGATTTTATATCATCTTGTCAAAATTCCTTCAGTTGTAAATGTCAAAAATGGCGCTGGGGA 48553
Qy 1090 AAGAGGTGGTGCCCTAAATGTTTAACTTGTAACTTGTCTTGTGCTCCCTGGGCACTT 1149
Db 48552 AAGAGGTGGTGCCCTAAATGTTTAACTTGTAACTTGTCTTGTGCTCCCTGGGCACTT 48493
Qy 1150 GGCTTTGTCTGCTCTCAGTGTCTTCCCTTTGACATGGGAAGAGTTGTGGCCAAAATC 1209
Db 48492 GGCTTTGTCTGCTCTCAGTGTCTTCCCTTTGACATGGGAAGAGTTGTGGCCAAAATC 48433
Qy 1210 CCCATCTTCTTGACCTCAACGCTCTGTGGCTCAGGGCTGGGGTGGCAGAGGAGGCGCTTC 1269
Db 48432 CCCATCTTCTTGACCTCAACGCTCTGTGGCTCAGGGCTGGGGTGGCAGAGGAGGCGCTTC 48373
Qy 1270 ACCTTATATCTGTTGTTTATCCAGGCTCCAGACTTCTCTCTGCTGCTCCCTCCCACTGCA 1329
Db 48372 ACCTTATATCTGTTGTTTATCCAGGCTCCAGACTTCTCTCTGCTGCTCCCTCCCACTGCA 48313
Qy 1330 CCCTCTCCCTTATCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1389
Db 48312 CCCTCTCCCTTATCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 48253
Qy 1390 CTCCTGGGTGATCCCTCCACTCTGACTCTGACTATGGCAGCAACACACAGGCGCTGGC 1449
Db 48252 CTCCTGGGTGATCCCTCCACTCTGACTCTGACTATGGCAGCAACACACAGGCGCTGGC 48193
Qy 1450 CCGAGTGATTTCAATGGTGATCAATTAATAAGAAAAATCGCAACAA 1495
Db 48192 CCGAGTGATTTCAATGGTGATCAATTAATAAGAAAAATCGCAACAA 48147

RESULT 9
AC110594
LOCUS
DEFINITION Homo sapiens chromosome 17, clone CTD-2193J24, complete sequence.
ACCESSION AC110594
VERSION AC110594.5 GI:22830329
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 159490)
Birren,B., Nusbaum,C. and Lander,E.
Unpublished
Homo sapiens chromosome 17, clone CTD-2193J24
REFERENCE
2 (bases 1 to 159490)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campoliano,A., Chang,J., Chazaro,B.,
Choepeil,Y., Colangelo,M., Collins,S., Collumore,A., Cook,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Lacroque,K., Lamazares,R.,
Landers,T., Lenockzy,J., Levine,R., Liu,G., Maclean,C.,
Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
```

```
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Triglio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (13-FEB-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 159490)
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepeil,Y., Collumore,A.,
Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (17-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 159490)
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepeil,Y., Collumore,A.,
Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (13-SEP-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 13, 2002 this sequence version replaced gi:22296745.
All repeats were identified using RepeatMasker:
Smith, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L25381
Center clone name: 2193_J_24
-----
Only the first 159,5 kilobases of this clone are being submitted.
The remainder overlaps accession number AC003042 [WICGR project
L260].
```


Db 151390 CCAGTGGATTTTCATGGTGATCATTTAAAAAGAAAAATCGCAACAA 151435

|||||

RESULT 10
AX834229 2547 bp DNA linear PAT 15-DEC-2003
LOCUS
DEFINITION Sequence 1353 from Patent EP1347046.
ACCESSION AX834229
VERSION AX834229.1 GI:39920364
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
1 Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R., Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahari, K. and Masuho, Y.

TITLE Full-length cDNA sequences
JOURNAL Patent: EP 1347046-A 1353 24-SEP-2003;
RESEARCH Association for Biotechnology (JP)

FEATURES
source
1..2547
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 42.8%; Score 644; DB 6; Length 2547;
Best Local Similarity 100.0%; Pred. No. 2.3e-113;
Matches 644; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 850 CAGATTGGAGATCCAGATGAGGCCACCGGAGCAGTGACCTAGTGGGAGCTC 909
Db 1904 CAGATTGGAGATCCAGATGAGGCCACCGGAGCAGTGACCTAGTGGGAGCTC 1963

QY 910 CTCCTTCCCTCCCACCTTTCAGTGGTCTGCTCCTGCTCTGATTTAGGTGTTGATT 969
Db 1964 CTCCTTCCCTCCCACCTTTCAGTGGTCTGCTCCTGCTCTGATTTAGGTGTTGATT 2023

QY 970 CTGGATCAGGGATACACTTCTGTCACACCCGACACGGGGCTAGAAAAATTTGTTG 1029
Db 2024 CTGGATCAGGGATACACTTCTGTCACACCCGACACGGGGCTAGAAAAATTTGTTG 2083

QY 1030 AGATTTTATATCATCTTGCAATGCTTCAGTTGTAATGTGAAAAATGGGCTGGGA 1089
Db 2084 AGATTTTATATCATCTTGCAATGCTTCAGTTGTAATGTGAAAAATGGGCTGGGA 2143

QY 1090 AAGGAGTGGTGTCCCTTAATGTTTACTTGTAACTTGTCTGTCGCCCTGGGCACTT 1149
Db 2144 AAGGAGTGGTGTCCCTTAATGTTTACTTGTAACTTGTCTGTCGCCCTGGGCACTT 2203

QY 1150 GGCTTTGTCTGCTCAGTGTCTTCCCTTGACATGGGAAGAGTTGTGGCCAAATC 1209
Db 2204 GGCTTTGTCTGCTCAGTGTCTTCCCTTGACATGGGAAGAGTTGTGGCCAAATC 2263

QY 1210 CCCATCTTCTTGACCTCAAGCTCTGGCTCGAGGCTCGGGTGGGAGAGGAGCCCTTC 1269
Db 2264 CCCATCTTCTTGACCTCAAGCTCTGGCTCGAGGCTCGGGTGGGAGAGGAGCCCTTC 2323

QY 1270 ACCATTATATCTGTTGTTATCCAGGCTCCAGACTTCTCTCTGCTGCTGCCCACTGCA 1329
Db 2324 ACCATTATATCTGTTGTTATCCAGGCTCCAGACTTCTCTCTGCTGCTGCCCACTGCA 2383

QY 1330 CCCTCTCCCCCTTATCATCTCTTCTGGCTCCCCAGCCAGTCTTGCTTCTTCTCCC 1389
Db 2384 CCCTCTCCCCCTTATCATCTCTTCTGGCTCCCCAGCCAGTCTTGCTTCTTCTCCC 2443

QY 1390 CTCCTGGGCTATCCCTCCACTCTGACTCTGACTATGCGACAGAAACACGAGGCCCTGGC 1449
Db 2444 CTCCTGGGCTATCCCTCCACTCTGACTCTGACTATGCGACAGAAACACGAGGCCCTGGC 2503

QY 1450 CCAGTGGATTTTCATGGTGATCATTTAAAAAGAAAAATCGCAACC 1493
Db 2504 CCAGTGGATTTTCATGGTGATCATTTAAAAAGAAAAATCGCAACC 2547

|||||

RESULT 11
AK096551 2547 bp mRNA linear PRI 15-JUL-2002
LOCUS
DEFINITION Homo sapiens cDNA FLJ39232 fis, clone OCBBF2007622.
ACCESSION AK096551
VERSION AK096551.1 GI:21756072
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
1 Oshima, A., Takahashi-Fujii, A., Tanase, T., Imose, N., Takeuchi, K., Arita, M., Musashino, K., Yuuki, H., Hara, H., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Ota, T., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T.

TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2547)
AUTHORS Isogai, T. and Yamamoto, J.
TITLE Direct Submission
JOURNAL Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)

COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES
source
1..2547
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="OCBPF2007622"
/tissue_type="brain"
/clone_lib="OCBPF2"
/dev_stage="fetal"
/note="cloning vector: pME18SFL3"

ORIGIN
Query Match 42.8%; Score 644; DB 9; Length 2547;
Best Local Similarity 100.0%; Pred. No. 2.3e-113;
Matches 644; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 850 CAGATTGGAGATCCAGATGAGGCCACCGGAGCAGTGACCTAGTGGGAGCTC 909
Db 1904 CAGATTGGAGATCCAGATGAGGCCACCGGAGCAGTGACCTAGTGGGAGCTC 1963

QY 910 CTCCTTCCCTCCCACCTTTCATGGCTTGCTCCTGCTCTGATTTAGGTGTTGATT 969
Db 1964 CTCCTTCCCTCCCACCTTTCATGGCTTGCTCCTGCTCTGATTTAGGTGTTGATT 2023

QY 970 CTGGATCAGGGATACCACTTCTGCTCCACACCCGACACGGGGCTAGAAAAATTTGTTG 1029
Db 2024 CTGGATCAGGGATACCACTTCTGCTCCACACCCGACACGGGGCTAGAAAAATTTGTTG 2083

QY 1030 AGATTTTATATCATCTTGCAATGCTTCAGTTGTAATGTGAAAAATGGGCTGGGA 1089
Db 2084 AGATTTTATATCATCTTGCAATGCTTCAGTTGTAATGTGAAAAATGGGCTGGGA 2143

QY 1090 AAGAGGTGGTGTCCCTAATGTTTACTTGTAACTGTTTGTGTCCTTGGCACTT 1149
 Db 2144 AAGAGGTGGTGTCCCTAATGTTTACTTGTAACTGTTTGTGTCCTTGGCACTT 2203
 QY 1150 GGCCTTGTCTGCTCTCAGTGTCTTCCCTTTGACATGGAAAGAGTTGTGGCCAAATC 1209
 Db 2204 GGCCTTGTCTGCTCTCAGTGTCTTCCCTTTGACATGGAAAGAGTTGTGGCCAAATC 2263
 QY 1210 CCCATCTTCTGCACTCAACGCTGTGGCTCAGGGCTGGGGTGGCAGAGGAGCCCTTC 1269
 Db 2264 CCCATCTTCTGCACTCAACGCTGTGGCTCAGGGCTGGGGTGGCAGAGGAGCCCTTC 2323
 QY 1270 ACCTTATATCTGTGTGTTATCCAGGGCTCCAGACTTCCCTCTGCTGCTGCCCACTGCA 1329
 Db 2324 ACCTTATATCTGTGTGTTATCCAGGGCTCCAGACTTCCCTCTGCTGCTGCCCACTGCA 2383
 QY 1330 CCCTCTCCCTTATCTATCTCTCTGCTGCTGCCCACTGCTTGTGCTTCTTGTCCC 1389
 Db 2384 CCCTCTCCCTTATCTATCTCTCTGCTGCTGCCCACTGCTTGTGCTTCTTGTCCC 2443
 QY 1390 CTCCTGGGTCATCCCTCCACTCTGACTCTGACTATGGCAGCAGAACACAGGGCCTTGGC 1449
 Db 2444 CTCCTGGGTCATCCCTCCACTCTGACTCTGACTATGGCAGCAGAACACAGGGCCTTGGC 2503
 QY -450 CCAGTGATTTCATGTGTGATCATTAATAAAGAAAAATCGCAACC 1493
 Db 2504 CCAGTGATTTCATGTGTGATCATTAATAAAGAAAAATCGCAACC 2547
 RESULT 12
 AX768992
 LOCUS AX768992 2077 bp DNA linear PAT 02-JUL-2003
 DEFINITION Sequence 109 from Patent WO02098917.
 ACCESSION AX768992
 VERSION AX768992.1 GI:32437160
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Guo, X., Fernandes, E., Li, L., Kekuda, R., Liu, Y., Leite, M.,
 Spytek, K.A., Ji, W., Casman, S.J., Boldog, F.L., Patturajan, M.,
 Vernet, C.A., Ballinger, R.A., Malyankar, U.M., Tchernev, V.T.,
 Blalock, A.D., Gusev, V.Y., Rastelli, L., Mezes, P.D., Ellerman, K.,
 Heyes, M., Herkmann, J.L., Shimkets, R.A., Toime, N., Pena, C.E.,
 Shenoy, S.G., Taupier, R.J., Gerlach, V. and Gorman, L.
 TITLE Human proteins and nucleic acids encoding same
 JOURNAL Patent: WO 02098917-A 109 12-DEC-2002;
 Curagen Corporation (US)
 FEATURES
 Location/Qualifiers
 1..2077
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Query Match 40.0%; Score 602; DB 6; Length 2077;
 Best Local Similarity 83.8%; Pred. No. 2.6e-105;
 Matches 759; Conservative 0; Mismatches 0; Indels 147; Gaps 1;
 QY 127 ATGAGCGGTGGCGACCGCTGGCGTGGTGTGACGGGGCTCGGGGGCATCGGCG 186
 Db 1 ATGAGCGGTGGCGACCGCTGGCGTGGTGTGACGGGGCTCGGGGGCATCGGCG 60
 QY 187 GCGGTGGCGCGCGCTGGTCCAGCAGGACTGAAGTGTGGTGGTGGCGCGCACTGTG 246
 Db 61 GCGGTGGCGCGCGCTGGTCCAGCAGGACTGAAGTGTGGTGGTGGCGCGCACTGTG 120
 QY 247 GGCACATCGAGGAGCTGGCTGCTGAATGTAAAGTGCAGGCTACCGCGGACTTGTATC 306
 Db 121 GGCACATCGAGGAGCTGGCTGCTGAATGTAAAGTGCAGGCTACCGCGGACTTGTATC 180

QY 307 CCTACAGATGTGACCTATCAATGAAGAGGACATCTCTCCATGTTCTCAGCTATCCGT 366
 Db 181 CCTACAGATGTGACCTATCAATGAAGAGGACATCTCTCCATGTTCTCAGCTATCCGT 240
 QY 367 TCTCAGCAGCGGTGTAGACATCTGCATCAACAATGTGGCTTGGCCCGGCTGACACC 426
 Db 241 TCTCAGCAGCGGTGTAGACATCTGCATCAACAATGTGGCTTGGCCCGGCTGACACC 300
 QY 427 CTGCTCTCAGCAGCAGCAGTGGTTGGAAGGACATGTTCAATGTGAAGTGTGGCCCTC 486
 Db 301 CTGCTCTCAGCAGCAGCAGTGGTTGGAAGGACATGTTCAATGTGAAGTGTGGCCCTC 360
 QY 487 AGCATCTGCACAGGGAGGCTTACAGTCCATGAAGAGGAGGAATGTGGACGATGGGAC 546
 Db 361 AGCATCTGCACAGGGAGGCTTACAGTCCATGAAGAGGAGGAATGTGGACGATGGGAC 420
 QY 547 ATCATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGTCTGACCCACTTC 606
 Db 421 ATCATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGTCTGACCCACTTC 480
 QY 607 TATAGTGCCACCAAGTATGCCGTCACTGCGCTGCACAGGAGGACTGAGGAGAGCTTCGG 666
 Db 481 TATAGTGCCACCAAGTATGCCGTCACTGCGCTGCACAGGAGGACTGAGGAGAGCTTCGG 540
 QY 667 GAGGCCAGACCCACATCCGAGCCAC----- 692
 Db 541 GAGGCCAGACCCACATCCGAGCCACCTCGGAGGAGGAGGAGGCGGCTGCCGGA 600
 QY 693 ----- 692
 Db 601 TATCAGGAGCCCATCACTGTGAAGCTGGGTTCTGTGGCTCCATCCTCTCCCTCGACC 660
 QY 693 -----GTGCATC 699
 Db 661 TCCCCAAGACCTGGCAAGCTCAGCCCTCAGAGGCCCTCTCTGTGGCCCATGTGCATC 720
 QY 700 TCTCCAGGTGTGGTGAGACACAATTCGCTTCAAACTCCACAGAGGAGGAGGAGGAG 759
 Db 721 TCTCCAGGTGTGGTGAGACACAATTCGCTTCAAACTCCACAGAGGAGGAGGAGGAG 780
 QY 760 CGAGCTGCCACCTATGAGCAAAATGAAGTGTCTCAAACTCCAGGAGTGTGGCCGAGGCTGT 819
 Db 781 CGAGCTGCCACCTATGAGCAAAATGAAGTGTCTCAAACTCCAGGAGTGTGGCCGAGGCTGT 840
 QY 820 ATCTAGTCTCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 879
 Db 841 ATCTAGTCTCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 900
 QY 880 GAGCAG 885
 Db 901 GAGCAG 906

RESULT 13
 AX184145
 LOCUS AX184145
 DEFINITION Sequence 89 from patent US 6342581.
 ACCESSION AR184145
 VERSION AR184145.1 GI:20228114
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 1 (bases 1 to 569)
 REFERENCE Rosen, C.A., Ruben, S.M., Olsen, H.S. and Ebner, R.
 AUTHORS Secreted protein HHPF03
 TITLE Patent: US 6342581-A 89 29-JAN-2002;
 JOURNAL Location/Qualifiers
 FEATURES
 source
 1..569
 /organism="unknown"
 /mol_type="unassigned DNA"
 ORIGIN

```
Query Match          34.1%; Score 513.6; DB 6; Length 569;
Best Local Similarity 99.4%; Pred. No. 2.5e-88;
Matches 535; Conservative 1; Mismatches 0; Indels 2; Gaps 2;

QY 968 TTCGGATACGGGATACCACTTCCTGTCCACACCCGACAGGGGCTAGAAAATTTGTT 1027
DB 13 TTCGGATACGGGATACCA-TCCTGTCTMACACCCCGACAGGGGCTAGAAAATTTGTT 71

QY 1028 TGAGATTTTATATCATCTTGTCAAAATGCTTCAGTTGTAAATGTGAAAATGGGCTGGG 1087
DB 72 TGAGATTTTATATCATCTTGTCAAAATGCTTCAGTTGTAAATGTGAAAATGGGCTGGG 131

QY 1088 GAAAGAGGTGGTGTCCCTAAATGCTTTTACTTTTAACTTCTTGTGCCCTGGGCAC 1147
DB 72 TGAGATTTTATATCATCTTGTCAAAATGCTTCAGTTGTAAATGTGAAAATGGGCTGGG 131

QY 1088 GAAAGAGGTGGTGTCCCTAAATGCTTTTACTTTTAACTTCTTGTGCCCTGGGCAC 1147
DB 132 GAAAGAGGTGGTGTCCCTAAATGCTTTTACTTTTAACTTCTTGTGCCCTGGGCAC 191

QY 1148 TTGCCCTTTGCTCTCTCAGTGTCTCCCTTTTGTCCCTTTGACATGGGAAAGGAGTTGTGGCCAAAA 1207
DB 192 TTGCCCTTTGCTCTCTCAGTGTCTCCCTTTGACATGGGAAAGGAGTTGTGGCCAAAA 251

QY 1208 TCCCATCTTTTTCACACTCTCAACGTCTGTGGCTCAGGGCTGGGTGGCAGAGGGAGGCT 1267
DB 252 TCCCATCTTTTTCACACTCTCAACGTCTGTGGCTCAGGGCTGGGTGGCAGAGGGAGGCT 311

QY 1268 TCACCTTATATCTGTGTGTGTATCCAGGGCTCCAGACTTCCTCTCTGCTGCCCTGCTG 1327
DB 312 TCACCTTATATCTGTGTGTGTATCCAGGGCTCCAGACTTCCTCTCTGCTGCCCTGCTG 371

QY 1328 CACCTCTCCCTCTTATCTATCTCTCTTCTCGGCTCCCGAGTCTTGTGCTTCTTGTGTC 1387
DB 372 CACCTCTCCCTCTTATCTATCTCTCTTCTCGGCTCCCGAGTCTTGTGCTTCTTGTGTC 431

QY 1388 CCCTCTGGGTCATCTCCCTCACTCTGACTCTGACTATGGCAGCAGAACACAGGGGCTG 1447
DB 432 CCCTCTGGGTCATCTCCCTCACTCTGACTCTGACTATGGCAGCAGAACACCA-GGCCTG 490

QY 1448 GCCCAGTGGATTTTCATGGTGTATCAATAAAGAAAAATCGAACCAAAAAA 1505
DB 491 GCCCAGTGGATTTTCATGGTGTATCAATAAAGAAAAATCGAACCAAAAAA 548

RESULT 14
BD192674          569 bp      DNA      linear      PAT 17-JUL-2003
LOCUS             123 human secreted proteins.
DEFINITION        BD192674.1 GI:33002413
ACCESSION         BD192674
VERSION           JP 2002513295-A/88.
KEYWORDS          Homo sapiens (human)
SOURCE            Homo sapiens
ORGANISM          Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 569)
Fischer, C.L., Rosen, C.A., Soppet, D.R., Ruben, S.M., Kyaw, H., Li, Y.,
Zeng, Z., Lafleur, D.W., Moore, P.A., Shi, Y., Ols, H.S., Ebner, R. and
Brewer, L.A.
123 human secreted proteins
Patent: JP 2002513295-A 88 08-MAY-2002;
HUMAN GENOME SCIENCES INC
OS Homo sapiens (human)
PN JP 2002513295-A/88
PD 08-MAY-2002
PF 07-JUL-1998 JP 1999508744
PR 08-JUL-1997 US 60/051926, 08-JUL-1997 US 60/051929 PR
PR 08-JUL-1997 US 60/052793, 08-JUL-1997 US 60/051925 PR
08-JUL-1997 US 60/051931, 08-JUL-1997 US 60/051932 PR
08-JUL-1997 US 60/052803, 08-JUL-1997 US 60/052732 PR
08-JUL-1997 US 60/051916, 08-JUL-1997 US 60/051930 PR
08-JUL-1997 US 60/051918, 08-JUL-1997 US 60/051920 PR
08-JUL-1997 US 60/051919, 08-JUL-1997 US 60/051928 PR
08-JUL-1997 US 60/052795, 08-JUL-1997 US 60/052733 PR
18-AUG-1997 US 60/055948, 18-AUG-1997 US 60/055722 PR
```

```
18-AUG-1997 US 60/055723, 18-AUG-1997 US 60/055949 PR
18-AUG-1997 US 60/055964, 18-AUG-1997 US 60/055947 PR
18-AUG-1997 US 60/055950, 18-AUG-1997 US 60/055953 PR
18-AUG-1997 US 60/056360, 18-AUG-1997 US 60/055984 PR
18-AUG-1997 US 60/055984, 18-AUG-1997 US 60/055954 PR
12-SEP-1997 US 60/058785, 12-SEP-1997 US 60/058664 PR
12-SEP-1997 US 60/058660, 12-SEP-1997 US 60/058661 PI
L. FISCHER, CRAIG A ROSEN, DANIEL R SOPPET, STEVEN M RUBEN, PI
KYAW,
PI YI LI, ZHIZHEN ZENG, DAVID W LAFLEUR, PAUL A MOORE, YANGGU SHI, PI
HENRIK S OLSEN,
PI REINHARD EBNER, LAURIE A BREWER
PC C07H21/04.C12N15/63
CC n equals a,t,g, or c
FH Key Location/Qualifiers
FT misc feature (i).
FEATURES
source 1..569
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match          34.1%; Score 513.6; DB 6; Length 569;
Best Local Similarity 99.4%; Pred. No. 2.5e-88;
Matches 535; Conservative 1; Mismatches 0; Indels 2; Gaps 2;

QY 968 TTCGGATACGGGATACCACTTCCTGTCCACACCCGACAGGGGCTAGAAAATTTGTT 1027
DB 13 TTCGGATACGGGATACCA-TCCTGTCTMACACCCCGACAGGGGCTAGAAAATTTGTT 71

QY 1028 TGAGATTTTATATCATCTTGTCAAAATGCTTCAGTTGTAAATGTGAAAATGGGCTGGG 1087
DB 72 TGAGATTTTATATCATCTTGTCAAAATGCTTCAGTTGTAAATGTGAAAATGGGCTGGG 131

QY 1088 GAAAGAGGTGGTGTCCCTAAATGCTTTTACTTTTAACTTCTTGTGCCCTGGGCAC 1147
DB 132 GAAAGAGGTGGTGTCCCTAAATGCTTTTACTTTTAACTTCTTGTGCCCTGGGCAC 191

QY 1148 TTGCCCTTTGCTCTCTCAGTGTCTCCCTTTTGTCCCTTTGACATGGGAAAGGAGTTGTGGCCAAAA 1207
DB 192 TTGCCCTTTGCTCTCTCAGTGTCTCCCTTTGACATGGGAAAGGAGTTGTGGCCAAAA 251

QY 1208 TCCCATCTTTTTCACACTCTCAACGTCTGTGGCTCAGGGCTGGGTGGCAGAGGGAGGCT 1267
DB 252 TCCCATCTTTTTCACACTCTCAACGTCTGTGGCTCAGGGCTGGGTGGCAGAGGGAGGCT 311

QY 1268 TCACCTTATATCTGTGTGTGTATCCAGGGCTCCAGACTTCCTCTCTGCTGCCCTGCTG 1327
DB 312 TCACCTTATATCTGTGTGTGTATCCAGGGCTCCAGACTTCCTCTCTGCTGCCCTGCTG 371

QY 1328 CACCTCTCCCTCTTATCTATCTCTCTTCTCGGCTCCCGAGTCTTGTGCTTCTTGTGTC 1387
DB 372 CACCTCTCCCTCTTATCTATCTCTCTTCTCGGCTCCCGAGTCTTGTGCTTCTTGTGTC 431

QY 1388 CCCTCTGGGTCATCTCCCTCACTCTGACTCTGACTATGGCAGCAGAACACAGGGGCTG 1447
DB 432 CCCTCTGGGTCATCTCCCTCACTCTGACTCTGACTATGGCAGCAGAACACCA-GGCCTG 490

QY 1448 GCCCAGTGGATTTTCATGGTGTATCAATAAAGAAAAATCGAACCAAAAAA 1505
DB 491 GCCCAGTGGATTTTCATGGTGTATCAATAAAGAAAAATCGAACCAAAAAA 548

RESULT 15
AC015919/c
LOCUS             Homo sapiens chromosome 17 clone CTD-2047D24 map 17, LOW-PASS
DEFINITION        AC015919
ACCESSION         AC015919
VERSION           HTG; HTGS_PHASE0.
KEYWORDS          Homo sapiens (human)
SOURCE            Homo sapiens (human)
```

ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
TITLE	1 (bases 1 to 169713)	11329
JOURNAL	Birren,B., Linton,L., Nusbaum,C. and Lander,E.	12181
REFERENCE	Unpublished	12280: gap of 100 bp
AUTHORS	2 (bases 1 to 169713)	12181
	Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,	12181: contig of 852 bp in length
	Baldwin,J., Barna,N., Beckerly,R., Boguslavskiy,L., Boukhgalter,B.,	12280: gap of 100 bp
	Brown,A., Casale,A., Collings,M., Collins,S., Collamore,A.,	13121: contig of 841 bp in length
	Cooke,P., DeArrellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,	13121: gap of 100 bp
	Freireira,P., FitzHugh,M., Forrest,C., Funke,R., Gage,D.,	13222: contig of 849 bp in length
	Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,	14071: gap of 100 bp
	Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,	14171
	Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,	15030: contig of 860 bp in length
	McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,	15031: gap of 100 bp
	Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,	15131
	Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,	15131: contig of 844 bp in length
	Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,	15974: gap of 100 bp
	Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,	16075
	Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.	16075: contig of 826 bp in length
	Direct Submission	17001
	Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome	17001: gap of 100 bp
	Research, 320 Charles Street, Cambridge, MA 02141, USA	17335: contig of 835 bp in length
	On Jul 13, 2000 this sequence version replaced gi:6446748.	17335: gap of 100 bp
	All repeats were identified using RepeatMasker:	17335: contig of 840 bp in length
	Smit, A.F.A. & Green, P. (1996-1997)	17335: gap of 100 bp
	http://ftp.genome.washington.edu/RM/RepeatMasker.html	17335: contig of 846 bp in length
	----- Genome Center	17335: gap of 100 bp
	Center: Whitehead Institute/ MIT Center for Genome Research	17335: contig of 846 bp in length
	Center code: WIGR	17335: gap of 100 bp
	Web site: http://www-seq.wi.mit.edu	17335: contig of 846 bp in length
	Contact: sequence_submissions@genome.wi.mit.edu	17335: gap of 100 bp
	----- Project Information	17335: contig of 840 bp in length
	Center project name: L855	17335: gap of 100 bp
	Center clone name: 2047_D_24	17335: contig of 826 bp in length
	-----	17335: gap of 100 bp
	* NOTE: This record contains 177 individual	17335: contig of 848 bp in length
	* sequencing reads that have not been assembled into	17335: gap of 100 bp
	* contigs. Runs of N are used to separate the reads	17335: contig of 829 bp in length
	* and the order in which they appear is completely	17335: gap of 100 bp
	* arbitrary. Low-pass sequence sampling is useful for	17335: contig of 844 bp in length
	* identifying clones that may be gene-rich and allows	17335: gap of 100 bp
	* overlap relationships among clones to be deduced.	17335: contig of 845 bp in length
	* However, it should not be assumed that this clone	17335: gap of 100 bp
	* will be sequenced to completion. In the event that	17335: contig of 862 bp in length
	* the record is updated, the accession number will	17335: gap of 100 bp
	* be preserved.	17335: contig of 857 bp in length
	1	17335: gap of 100 bp
	841: contig of 841 bp in length	17335: contig of 843 bp in length
	941: gap of 100 bp	17335: gap of 100 bp
	942	17335: contig of 847 bp in length
	1778: contig of 837 bp in length	17335: gap of 100 bp
	1779	17335: contig of 847 bp in length
	1879	17335: gap of 100 bp
	2729: contig of 851 bp in length	17335: contig of 847 bp in length
	2730	17335: gap of 100 bp
	2829: gap of 100 bp	17335: contig of 858 bp in length
	2830	17335: gap of 100 bp
	3681: contig of 852 bp in length	17335: contig of 846 bp in length
	3682	17335: gap of 100 bp
	3781: gap of 100 bp	17335: contig of 846 bp in length
	3782	17335: gap of 100 bp
	4620: contig of 839 bp in length	17335: contig of 846 bp in length
	4621	17335: gap of 100 bp
	4720: gap of 100 bp	17335: contig of 843 bp in length
	4721	17335: gap of 100 bp
	5544: contig of 824 bp in length	17335: contig of 847 bp in length
	5545	17335: gap of 100 bp
	5644: gap of 100 bp	17335: contig of 847 bp in length
	6505: contig of 861 bp in length	17335: gap of 100 bp
	6506	17335: contig of 847 bp in length
	6606: gap of 100 bp	17335: gap of 100 bp
	6606	17335: contig of 842 bp in length
	7451: contig of 846 bp in length	17335: gap of 100 bp
	7452	17335: contig of 831 bp in length
	7551: gap of 100 bp	17335: gap of 100 bp
	8385: contig of 834 bp in length	17335: contig of 844 bp in length
	8485: gap of 100 bp	17335: gap of 100 bp
	8486	17335: contig of 833 bp in length
	9315: contig of 830 bp in length	17335: gap of 100 bp
	9316	17335: contig of 842 bp in length
	9415: gap of 100 bp	17335: gap of 100 bp
	10262: contig of 847 bp in length	17335: contig of 855 bp in length
	10263	17335: gap of 100 bp
	10362: gap of 100 bp	17335: contig of 834 bp in length
	10363	17335: gap of 100 bp
	11228: contig of 866 bp in length	17335: contig of 862 bp in length
	11229	17335: gap of 100 bp

2/26/20

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 16, 2004, 16:48:58 ; Search time 2446 Seconds
(without alignments)
3174.230 Million cell updates/sec

Title: US-10-063-735-128

Perfect score: 1337

Sequence: 1 MARGMERDRRLALVTGAS.....STPAHQIGDIQMPTEQVT 260

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 segs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlpl
-Q=/cgn2_1/USPTO_spool_p/US10063735/runat_15092004_103034_8933/app_query.fasta_1.455
-DB=EST -QPMT=fastcap -SUFFIX=rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -NATRX=blosum62 -TRANS=human40.cdi -List=100
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=50 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HSPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10063735@cgn_1_5180 @runat_15092004_103034_8933 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_in:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1138.5	89.6	829	12	BG324477 602422421
2	1178	88.1	822	12	BG741589 602635416
3	1117.5	83.6	1010	12	BM471232 AGENCOURT
4	1103.5	82.5	1071	13	BQ278156 AGENCOURT
5	1066	79.7	651	9	AV654573 AV654573
6	1061	79.4	897	13	BX777216 BX777216
7	1054.5	78.9	940	12	B1818880 603037354
8	1049.5	78.5	914	12	BG967186 602833867
9	1037	77.6	937	13	BUS39642 AGENCOURT
10	1025	76.7	932	10	BG030249 602297547
11	995.5	74.5	1063	10	BE796469 601589817
12	983	73.5	612	13	BX671560 BX671560
13	961	71.9	807	14	CK019037 AGENCOURT
14	956	71.5	769	14	CF221812 AGENCOURT
15	955	71.4	923	10	BE617298 601441804
16	948.5	70.9	852	12	BG282361 602402863
17	946	70.8	846	14	CA474188 AGENCOURT
18	939	70.2	947	14	CA964199 CcLL03a10
19	935	69.9	819	14	CA969358 CcLL06a22
20	934.5	69.9	881	10	BF233600 602024015
21	927	69.3	565	12	BM744024 K-EST0017
22	927	69.3	573	12	BM742038 K-EST0014
23	927	69.3	602	12	BM764342 K-EST0045
24	927	69.3	640	12	BM755987 K-EST0034
25	927	69.3	678	12	BI391549 p3p1n.pk0
26	919	68.7	849	13	BUI21233 603950353
27	912	68.2	743	13	EX864140 BX864140
28	908	67.9	912	14	CF595166 AGENCOURT
29	905	67.7	627	12	BM848012 K-EST0127
30	904	67.6	580	12	BM737591 K-EST0000
31	900	67.3	552	12	BM743569 K-EST0016
32	885.5	66.2	757	12	BI464353 603204005
33	883	66.0	1143	14	CF662388 CcLL01a04
34	881.5	65.9	714	12	BI546418 603188719
35	879	65.7	654	9	AL868596 AL868596
36	871	65.1	1001	14	CF225840 AGENCOURT
37	867	64.8	804	12	BI765897 603046146
38	851	63.6	935	13	BX773551 BX773551
39	846	63.3	946	10	BF581815 602099816
40	841.5	62.9	748	12	BI761329 603044093
41	837	62.6	974	14	CA966917 CcLL10a35
42	832	62.2	646	9	AL845993 AL845993
43	824	61.6	672	14	CD469548 LeukOS2.4
44	810	60.6	614	10	BE869917 601446563
45	807.5	60.4	977	13	BUS39980 AGENCOURT
46	801	59.9	567	14	W58472 zd25b06.r1
47	796.5	59.6	763	14	CF348186 AGENCOURT
48	792	59.2	648	9	AL898132 AL898132
49	786	58.8	650	14	CD599740 RK119A4F0
50	785	58.7	863	14	CA474613 AGENCOURT
51	783	58.6	607	12	BJ624590 BJ624590
52	782	58.5	455	10	BE395404 601309913
53	764	57.1	910	14	CF225841 AGENCOURT
54	763	57.1	650	9	AL895715 AL895715
55	759	56.8	507	14	W58514 zd25h06.r1
56	748	55.9	545	10	BF076147 225644.MA
57	748	55.9	623	12	BJ005084 BJ005084
58	747	55.9	634	12	BJ528471 BJ528471
59	745	55.7	540	12	BM848396 BM848396
60	745	55.7	751	10	BG037061 602288368
61	737	55.1	736	14	CD494676 CD412-B06
62	734	54.9	622	12	BJ502141 BJ502141

```
63 734 54.9 632 12 BJ497197
64 732 54.7 654 12 BJ013036
65 731 54.7 633 12 BJ489602
66 730 54.6 609 9 AL646156
67 729.5 54.6 890 14 CA970350
68 723 54.1 704 12 BJ518187
69 718 53.7 809 13 BJ090978
70 713 46.1 12 BJ088288
71 709 53.0 580 14 CA779637
72 707 52.9 628 14 CD496098
73 702 52.9 886 10 BE895089
74 699 52.3 548 12 BM316312
75 695 52.0 755 12 BM187418
76 694 51.9 571 12 BJ521969
77 676 50.6 696 10 BE395162
78 667 49.9 600 9 AL679337
79 653 48.8 547 10 AW640188
80 649 48.5 507 9 AI930229
81 633 47.3 531 9 AI545791
82 633 47.3 639 9 AL877407
83 632 47.3 606 14 CA329988
84 626 46.8 718 13 BU121297
85 624 46.7 516 12 BM708058
86 624 46.7 533 12 BJ014213
87 624 46.7 649 9 AL846856
88 622 46.5 649 12 BJ019121
89 619 46.3 837 13 BM057018
90 618 46.2 453 14 CD734701
91 616 46.1 554 14 CF788313
92 608 45.5 574 14 BF089313
93 607 45.4 663 13 BM307486
94 600 44.9 653 13 BM312777
95 599.5 44.8 1287 13 BU539488
96 598 44.7 871 13 BM128104
97 595 44.5 682 12 BU028379
98 590 44.1 721 12 BJ027323
99 581 43.5 750 13 BM181907
100 577 43.2 742 13 BM179752

ALIGNMENTS
RESULT 1
BG324477 829 bp mRNA linear EST 27-FEB-2001
LOCUS 602424242F1 NIH_MGC_14 Homo sapiens CDNA clone IMAGE:4560485 5',
DEFINITION mRNA sequence.
ACCESSION BG324477
VERSION BG324477.1 GI:13130914
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/BTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1269 row: j column: 06
High quality sequence stop: 755.
FEATURES
Location/Qualifiers
1..829
source
/organism="Homo sapiens"
/mol_type="mRNA"
```

```
/db_xref="taxon:9606"
/clone="IMAGE:4560485"
/tissue_type="renal cell adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 14"
/note="Organ: kidney; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adapter: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

ORIGIN

Alignment Scores:

Pred. No.:	9.29e-128	Length:	829
Score:	1198.50	Matches:	245
Percent Similarity:	96.86%	Conservative:	2
Best Local Similarity:	96.08%	Mismatches:	7
Query Match:	89.64%	Indels:	4
DB:	12	Gaps:	0

US-10-063-735-128 (1-260) x BG324477 (1-829)

QY	1	MetAlaArgProGlyMetGluArgTrrArgAspArgLeuAlaLeuValThrGlyAlaSer	20
DB	68	ATGGCCAGGCGCGGATGGAGCGGTGGCGCCGCGCTGGTACCGGGGCCCTCG	127
QY	21	GlyGlyLeuGlyAlaAlaValAlaArgAlaLeuValGlnGlnGlyLeuLysValValGly	40
DB	128	GGGGGATCGGGCGCGCGTGGCCCGGCCCTGGTCCAGCAGGAGCTGAAGGTGGTGGGC	187
QY	41	CysAlaArgThrValGlyAsnIleGluLeuAlaAlaGluCysLysSerAlaGlyTyr	60
DB	188	TGCGCCCGCAGTGTGGGCAACATCAGAGAGCTGGCTGTAATGTAAGAGTGCAGGCTAC	247
QY	61	ProGlyThrLeuIleProTyrArgCysAspLeuSerAsnGluAspIleLeuSerMet	80
DB	248	CCCGGGACTTTGATCCCCCTACAGATGTGACCTATCAATGAAGAGGAGCATCTCTCCATG	307
QY	81	PheSerAlaIleArgSerGlnHisSerGlyValAspIleCysIleAsnAsnAlaGlyLeu	100
DB	308	TTCTCAGCTATCCGTTCTTCAGCACACGCGGTGTAGACATCTGCATCAACAATGCTGCTTG	367
QY	101	AlaArgProAspThrLeuLeuSerGlySerThrSerGlyTrrLysAspMetPheAsnVal	120
DB	368	CCCCGGCCTGACACCCCTGCTCTCAGGCAGCACCCAGTGGTGGAAAGGACATGTTCAATGTG	427
QY	121	AsnValLeuAlaLeuSerIleCysThrArgGluAlaTyrGlnSerMetLysGluArgAsn	140
DB	428	AACGTGTGGCCCTCAGCATCTGCACCGGAACCTACCGATCCATGAAGAGCGGAAT	487
QY	141	ValAspAspGlyHisIleIleAsnIleAsnSerMetSerGlyHisArgValLeuProLeu	160
DB	488	GTGGACGATGGGCACATCATTAACATCAATAGCATGCTGCGCCACCGAGTGTACCCCTG	547
QY	161	SerValThrHisPheTyrSerAlaThrLysTyrAlaValThrAlaLeuThrGluGlyLeu	180
DB	548	TCTGTGACCCACTTCTATAGTCCCAAGTATGCCGCTCACTGGCGCTGACAGAGGACTG	607
QY	181	ArgGlnGluLeuArgGluAlaGlnThrHisIleArgAlaThrCysIleSerProGlyVal	200
DB	608	AGGCAAGAGCTTCGGAGGCCACACCCATCCAGCCAGCTGTCATCTCTCCAGGTGTG	667
QY	201	ValGluThr-GlnPheAlaPheLysLeuHisAspLysAspProGlyLysAlaAlaAlaTh	220
DB	668	GTGGAGACCAAAATTCGCTTCAAACTCCACGACCAAGGACCTTGAGAGGAGCATCTGCANC	727
QY	220	rTyrGluGlnMetLysCysLeuLysProGluAspValAlaGluAlaValIleTyrValLe	240
DB	728	-TATGAGCAATGAAGTGTCTCAAAACCCGAGGATGTGGCGGAGGCTGTATCTACGT-CT	785


```

Qy      240 userThrProalaHisIleGlnIleGlyAspIleGlnMetArg 254
      ||||||| ||||| ||||||| ||||||| ||||| |||||
Db      786 CAGCACCCCGGA-CACATTCGATGGGAGCTCCGATGAAGG 827

RESULT 2
LOCUS   BG741589                               822 bp   mRNA   linear   EST 15-MAY-2001
DEFINITION 602635416F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4780659 5',
      mRNA sequence.
ACCESSION BG741589
VERSION   BG741589.1 GI:14052242
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1. (bases 1 to 822)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL   National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT   Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10639 row: p column: 04
High quality sequence stop: 803.
FEATURES             Location/Qualifiers
     source
     1..822
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="IMAGE:4780659"
     /lab_host="DH10B (T1 phage-resistant)"
     /clone_lib="NCI_CGAP_Skn3"
     /note="organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
     Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
     Average insert size 1.5kb. Library constructed by Life
     Technologies. Note: this is a NCI_CGAP Library."

ORIGIN
Alignment Scores:
Pred. No.:      2,17e-125      Length:      822
Score:          1178.00      Matches:    234
Percent Similarity: 98.32%      Conservative: 0
Best Local Similarity: 98.32%      Mismatches:  2
Query Match:    88.11%      Indels:     2
DB:             12           Gaps:       0

US-10-063-735-128 (1-260) x BG741589 (1-822)

Qy      1 MetAlaArgProGlyMetGluArgTrpArgAspArgLeuAlaLeuValThrGlyAlaSer 20
      ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db      102 ATGGCCAGGCCCGGATGGAGCGGTGGCGCGACCGGCTGGCGCTGGTGCAGGGGGCCCTCG 161

Qy      21 GlyGlyIleGlyAlaAlaValAlaArgAlaLeuValGlnGlnGlyLeuLysValValGly 40
      ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db      162 GGGGGCATCGGCGCGCGCTGGCCGGGCCCTGGTCCAGCAGGAGCTGNAGGTGTGGGC 221

Qy      41 CysAlaArgThrValGlyAsnIleGluGluLeuAlaAlaGluCysGlySerAlaGlyTyr 60
      ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db      222 TCGGCCCGCACGTGTGGGCAACATCGAGGAGCTGGCTGCTGAATGTAAAGAGTCAGGCTAC 281

Qy      61 ProGlyThrLeuIleProTyrArgCysAspLeuSerAsnGluGluAspIleLeuSerMet 80
      ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db      282 CCCGGAGCTTTCATCCCTACAGATGTGACCTATCAAAATGAAGAGGACATCCTCTCCATG 341

Qy      81 PheSerAlaIleArgSerGlnHisSerGlyValAspIleCysIleAsnAlaGlyLeu 100
      ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

```

```

Db      342 TTCTCAGCTATCCGTTCTCAGCACAGCGGTGTAGACATCTGCATCAACAATGCTGGCTTG 401
Qy      101 AlaArgProAspThrLeuLeuSerGlySerThrSerGlyTrpLysAspMetPheAsnVal 120
      ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db      402 GCCCGGCTGACACCCCTGCTCTCAGCGCACACCACTGCTGGTGTGGAAGGACATGTTCATATGTG 461
Qy      121 AsnValLeuAlaLeuSerIleCysThrArgGluAlaTyrGlnSerMetLysGluArgAsn 140
      ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db      462 AACGTGCTGGCCCTCAGCATCTGCACAGGAAGCTACCACTCCATGAGGAGCGGAAT 521
Qy      141 ValAspAspGlyHisIleIleAsnIleAsnSerMetSerGlyHisArgValLeuProLeu 160
      ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db      522 GTGGACGATGGCACATCATTAACATCAATAGCATGTCTGCCACCGAGTGTATCCCTGTG 581
Qy      161 SerValThrHisPheTyrSerAlaThrLysTyrAlaValThrAlaLeuThrGluGlyLeu 180
      ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db      582 TCTGTGACCCACTTCTATAGTGCACCAAGTATGCCGTCACTGCCTGACAGGAGGACTG 641
Qy      181 ArgGlnGluLeuArgGluAlaGlnThrHisIleArgAlaThrCysIleSerProGlyVal 200
      ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db      642 AGGCAAGAGCTTCGGGAGGCCACACCATCCGAGCCAGTGCATCTCCAGGTGTG 701
Qy      201 ValGluThrGlnPheAlaPheLysLeuHisAspLysAspProGluLysAlaAlaAlaThr 220
      ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db      702 GTGGAGACACAAATTCGACTTCAAACCTCCACGACAGGACCTGAGAGGAGTGCCACC 761
Qy      221 Tyr-GluGlnMetLys-CysLeuLysProGluAspValAlaGluAlaVal 236
      ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db      762 TATTGAGCAATGAACGTGTCTCAAAACCGAGGATGTGGCCGAGGCTGTA 811

RESULT 3
LOCUS   BM471232                               1010 bp   mRNA   linear   EST 05-FEB-2002
DEFINITION 6478395 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5563124
      5', mRNA sequence.
ACCESSION BM471232
VERSION   BM471232.1 GI:18520274
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1. (bases 1 to 1010)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL   National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT   Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12293 row: j column: 21
High quality sequence stop: 561.
FEATURES             Location/Qualifiers
     source
     1..1010
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="IMAGE:5563124"
     /tissue_type="melanotic melanoma"
     /lab_host="DH10B (phage-resistant)"
     /clone_lib="NIH_MGC_72"
     /note="organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
     Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
     Average insert size 2 kb. Library constructed by Life
     Technologies."

ORIGIN
Alignment Scores:

```

```

Pred. No.: 3,15e-118 Length: 1010
Score: 1117.50 Matches: 235
Percent Similarity: 91.70% Conservative: 8
Best Local Similarity: 88.68% Mismatches: 15
Query Match: 83.58% Indels: 7
DB: 12 Gaps: 1

US-10-063-735-128 (1-260) x BM471232 (1-1010)

QY 1 MetAlaArgProGlyMetGluArgTrpArgAspArgLeuAlaLeuValThrGlyAlaSer 20
Db 87 ATGGCCAGCCCGGATGAGCGGTGGCGCGACCGCGTGGCGTGGTGAACGGGGCTCG 146
QY 21 GlyGlyIleGlyAlaAlaValAlaAlaAlaLeuValGlnGlnGlyLeuValValGly 40
Db 147 GGGGGCATCGGGCGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 206
QY 41 CysAlaArgThrValGlyAsnIleGluGluLeuAlaAlaGluCysLysSerAlaGlyTyr 60
Db 207 TCGCGCCGACCTGTGGCAACATCGAGGAGCTGGCTGTAATGTAAGATGTCAGGCTAC 266
QY 61 ProGlyThrLeuIleProTyrArgCysAspLeuSerAsnGluGluAspIleLeuSerMet 80
Db 267 CCGGGGACTTTGATCCCTACAGATGTGACCTATCAATGAAGAGGACATCTCTCCATG 326
QY 81 PheSerAlaIleArgSerGlnHisSerGlyValAspIleCysIleAsnAsnAlaGlyLeu 100
Db 327 TTCTCAGCTATCCGTCTCAGCACAGCGGTGTAGACATCTGATCAACATGCTGGCTTG 386
QY 101 AlaArgProAspThrLeuLeuSerGlySerThrSerGlyTyrIleAspMetPheAsnVal 120
Db 387 GCGCGGCTGACACCTGCTCTCAGGCAGCAGCAGCGTGGTGAAGGACATGTTCAATG 446
QY 121 AsnValLeuAlaLeuSerIleCysThrArgGluAlaTyrGlnSerMetLysGluArgAsn 140
Db 447 AACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCTACCATGTCATGAAGGAGGGAAT 506
QY 141 ValAspAspGlyHisIleIleAsnIleAsnSerMetSerGlyHisArgValIleuProLeu 160
Db 507 GTGGACGATGGGCATCATATTAACTATCAATGATGCTGGCCACCGAGTGTACCCCTG 566
QY 161 SerValThrHisPheTyrSerAlaThrIleValAlaValThrAlaLeuThrGluGlyLeu 180
Db 567 TCTGTGACCCACTTCTATAGTCCACCAAGTATGCGTCACTGCGTGCAGACGGGACTG 626
QY 181 ArgGlnGluLeuArgGluAlaGlnThrHisIleArgAlaThrCysIleSerProGlyVal 200
Db 627 AGGCAAGAGCTTCGGGAGGCCAGACCCACATCCGAGCCACGTCATCTCTCAGGGGG 686
QY 201 ValGluThrGlnPheAlaPheLysLeuHisAspLysAspPro-GluLysAlaAlaA-T 220
Db 687 GTGGAGACACCATTCGGCTTCAAACTCCCGACCAAGACCTTGGAGAGGAGCTGCCCC 746
QY 220 hrTyrGluGlnMetLysCys-LeuLys-ProGluAspValAlaGluAlaValIleTyr-V 239
Db 747 CCTATGGACCAATGAAGTGTCTCCAAACCCCAAGATGGGCCAAGGCTGTATCTACGG 806
QY 239 aIleuSerThrPro---AlaHisIleGlnIleGlyAspIleGlnMetArg-ProThrGlu 257
Db 807 TCTCAGGAACCCCGAACACATTCAGATTGGAAATACTCCCAAAATGAGGGCCCGGAA 866
QY 258 Gln 258
Db 367 CAG 869

RESULT 4
BQ278156
LOCUS
DEFINITION BQ278156 1071 bp mRNA linear EST 07-MAY-2002
5' mRNA sequence.
ACCESSION BQ278156
VERSION BQ278156.1 GI:20488364
KEYWORDS EST.

```

```

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1071)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM2043 row: n column: 09
High quality sequence stop: 603.
FEATURES
    source
        1..1071
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5805128"
            /tissue_type="teratocarcinoma, cell line"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH MGC_109"
            /note="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2:
            XhoI; cDNA made by oligo-dT priming. Directionally cloned
            into EcoRI/XhoI sites using the following 5' adaptor:
            GGCAGCAG(G). Library constructed by Ling Hong in the
            laboratory of Gerald M. Rubin (University of California,
            Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
            Superscript II RT (Life Technologies). Note: this is a
            NIH_MGC Library."
ORIGIN
Alignment Scores:
Pred. NO.: 1,46e-116 Length: 1071
Score: 1103.50 Matches: 230
Percent Similarity: 93.25% Conservative: 5
Best Local Similarity: 91.27% Mismatches: 10
Query Match: 82.54% Indels: 7
DB: 13 Gaps: 3

US-10-063-735-128 (1-260) x BQ278156 (1-1071)

QY 1 MetAlaArgProGlyMetGluArgTrpArgAspArgLeuAlaLeuValThrGlyAlaSer 20
Db 96 ATGGCCAGCCCGGATGAGCGGTGGCGCGACCGCGTGGCGTGGTGAACGGGGCTCG 155
QY 21 GlyGlyIleGlyAlaAlaValAlaAlaAlaLeuValGlnGlnGlyLeuLysValValGly 40
Db 156 GGGGGCATCGCGCGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 215
QY 41 CysAlaArgThrValGlyAsnIleGluGluLeuAlaAlaGluCysLysSerIleGlyTyr 60
Db 216 TCGCGCCGACCTGTGGCAACATCGAGGAGCTGGCTGTAATGTAAGATGTCAGGCTAC 275
QY 61 ProGlyThrLeuIleProTyrArgCysAspLeuSerAsnGluGluAspIleLeuSerMet 80
Db 276 CCGGGGACTTTGATCCCTACAGATGTGACCTATCAATGAAGAGGACATCTCTCCATG 335
QY 81 PheSerAlaIleArgSerGlnHisSerGlyValAspIleCysIleAsnAsnAlaGlyLeu 100
Db 336 TTCTCAGCTATCCGTCTCAGCACAGCGGTGTAGACATCTGCATCAACATGCTGGCTTG 395
QY 101 AlaArgProAspThrLeuLeuSerGlySerThrSerGlyTyrIleAspMetPheAsnVal 120
Db 396 GCGCGGCTGACACCCCTCTCTCAGGAGCAGCAGCTGGTGGTGAAGGACATGTTCAATGTG 455
QY 121 AsnValLeuAlaLeuSerIleCysThrArgGluAlaTyrGlnSerMetLysGluArgAsn 140

```

```

Db      |||||||AACGTCGTGGCCTCAGCATCTGCACAGGAGCCTACAGTCATGAGGAGCGGAAT 515
Qy      |||||||ValAspAspGlyHisIleAsnIleAsnSerMetSerGlyHisArgValLeuProLeu 160
Db      |||||||516 GTGGACGATGGGCACATCATTAACATCAATAGATGTCTGGCCACCGAGTGTACCCCTG 575
Qy      |||||||161 SerValThrHisPheTyrSerAlaThrIleValThrAlaLeuThrGluGlyLeu 180
Db      |||||||576 TCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTCACTGCCCTGACAGGAGCTG 635
Qy      |||||||181 ArgGlnGluLeuArgGlu-AlaGlnThrHisIleArgAlaThrCysIleSerProGlyVa 200
Db      |||||||636 AGGCAAGAGCTTCGGAGGGGCCAGACCCACATCCGAGCCAGTCATCTCTCCAGTGT 695
Qy      |||||||200 1-ValGluThrGlnPheLeuPheLeuHisAspLysAspPro---GluLysAlaAla 219
Db      |||||||696 GGGTGGAGACAAATCCCTTCAAACTCCACGACGAGGACCTCGAGAGGGCACCTGG 755
Qy      |||||||219 laThrTyrGluGlnMetLysCys-LeuLysProGluAsp-ValAlaGluAla---ValI 237
Db      |||||||756 CCCCCTATGAACAAATGAAGTGTCTCAAAACCAAGGAATGTGGCCCGAAGCTGTTATC 815
Qy      |||||||237 eTyrValLeu---SerThrProAlaHis 245
Db      |||||||816 TTACGTCTTTCAGCAACCCCGGCACAC 843

```

RESULT 5
 AV654573 651 bp mRNA linear EST 15-JAN-2002
 LOCUS AV654573 GLC Homo sapiens cDNA clone GLCDXC05 3', mRNA sequence.
 DEFINITION AV654573
 ACCESSION AV654573
 VERSION AV654573.1 GI:9875587
 KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 651)
 Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
 Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
 Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
 Hu G., Gu,J., Chen,Z., and Han,Z.
 Insight into hepatocellular carcinogenesis at transcriptome level
 by comparing gene expression profiles of hepatocellular carcinoma
 with those of corresponding noncancerous liver
 Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
 21625106
 11752456
 Contact: Zeguang Han
 Chinese National Human Genome Center at Shanghai
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
 201203, P. R. China
 Tel: 86-21-50801919 (ex. 45)
 Fax: 86-21-50801922
 Email: hanzg@chgc.sh.cn
 This clone is available at CHGC in Shanghai.

FEATURES
 source
 1..651
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="GLCDXC05"
 /tissue_type="corresponding non cancerous liver tissue"
 /dev_stage="Adult"
 /lab_host="SOLR"
 /clone_lib="GLC"
 /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
 XhoI"

Alignment Scores: 1.34e-112 Length: 651

ORIGIN

Pred. No.: 1.34e-112 Length: 651

Score: 1066.00 Matches: 208
 Percent Similarity: 97.22% Conservative: 2
 Best Local Similarity: 96.30% Mismatches: 6
 Query Match: 79.73% Indels: 0
 DB: 9 Gaps: 0

US-10-063-735-128 (1-260) x AV654573 (1-651)

Qy 11 AspArgLeuAlaLeuValThrGlyAlaSerGlyGlyIleGlyAlaAlaValAlaArgAla 30
 Db 2 GACCGGCTGGCGCTGGTGCACGGGGGCGCTCGGGGGGCATCGCGCGCGCTGCCCGGGCC 61
 Qy 31 LeuValGlnGlnGlyValGlyValGlyCysAlaArgThrValGlyAsnIleGluGlu 50
 Db 62 CTGGTCCAGCAGGAGACTGAAGGTGGTGGGCTGCGCCCGCACTGTGGGCAATCATCGAGG 121
 Qy 51 LeuAlaAlaGluCysLysSerAlaGlyTyrProGlyThrLeuIleProTyrArgCysAsp 70
 Db 122 CTGGCTGCTGAATGTAGAGTGCAGGCTACCCCGGACTTTGATCCCTTACAGATGTGAC 181
 Qy 71 LeuSerAsnGluGluAspIleLeuSerMetPheSerAlaIleArgSerGlnHisSerGly 90
 Db 182 CTATCAATGAAGAGGACATCTCTCCATGTCTCAGCTATCCGTTCTCAGCACACGCGT 241
 Qy 91 ValAspIleCysIleAsnAsnAlaGlyLeuAlaArgProAspThrLeuLeuSerGlySer 110
 Db 242 GTAGACATCTGCATCAACAATGCTGGCTGGCCCGGCTGACACCTCTCTCAGGCGAGC 301
 Qy 111 ThrSerGlyTyrLysAspMetPheAsnValAsnValLeuAlaLeuSerIleCysThrArg 130
 Db 302 ACCAGTGGTGGAGGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACACGG 361
 Qy 131 GluAlaTyrGlnSerMetLysGluArgAsnValAspAspGlyHisIleIleAsnIleAsn 150
 Db 362 GAAGCCTACCACTGCATGAAGGAGCGAATGTGGACGATGGCGACATCATTAACATCAAT 421
 Qy 151 SerMetSerGlyHisArgValLeuProLeuSerValThrHisPheTyrSerAlaThrLys 170
 Db 422 AGCATGTCTGGCCACCGAGTGTATCCCTGCTGTGACCCCACTTCTATAGTGCCACCAAG 481
 Qy 171 TyrAlaValThrAlaLeuThrGluGlyLeuArgGlnGluLeuArgGluAlaGlnThrHis 190
 Db 482 TATGCCGTCACTGGCTGCACAGAGGACTGAGGCAGAGCTTCGAGGCGCCAGACCCAC 541
 Qy 191 IleArgAlaThrCysIleSerProGlyValValGlnThrGlnPheAlaPheLysLeuHis 210
 Db 542 ATCCGAGCAGCTGCATCTCTTCAGGTGTGGGTGAGACACAATATATCGCTTTAAACTTCAC 601
 Qy 211 AspLysAspProGluLysAlaAlaAlaThrTyrGluGlnMetLysCys 226
 Db 602 GACAGGGACCTTGAGAGGCGACTGGCAACCTATGAGCAAGGAAGTGC 649

RESULT 6

LOCUS BX777216 897 bp mRNA linear EST 10-DEC-2003
 DEFINITION BX777216 XGC-egg Silurana tropicalis cDNA clone TEGg062g24 3', mRNA
 sequence.
 ACCESSION BX777216
 VERSION BX777216.1 GI:39684422
 KEYWORDS EST.
 SOURCE Silurana tropicalis (western clawed frog)

ORGANISM

Silurana tropicalis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 Xenopodinae; Silurana.

1 (bases 1 to 897)

REFERENCE
 Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
 Sanger Xenopus tropicalis EST project 2001 (11_2003)
 TITLE Unpublished (2003)
 JOURNAL
 COMMENT Contact: Croning MDR
 Sanger Institute
 Hinxton, Cambridgeshire, CB10 1SA, UK
 Email: trop@sanger.ac.uk


```

Db      127 ATGGCAGGCCCGCATGAGCGGTGGCGCGACCGGCTGGCGCTGTGTGTGACGGGGCCCTC 186
QY      20 rGly---GlyIleGlyAlaAlaValAlaArg-AlaLeuValGlnGlnGlnGlyLeuLysValY 39
Db      187 GGGGTGGCATCTGTGGTCGGCGCGTGGCCCTGGGCCCTGTGTCCAGCAGGAGCTGAAGGTGG 246
QY      39 aIdGlyCysAlaArgThrValGlyAsnIleGluGluLeuAlaGluCysLysSerAlaG 59
Db      247 TGGGCTGCCCGCATCTGTGGCAACATCGAGGAGTGGCTGCTGAATGTAGAGTGCAG 306
QY      59 lYrProGlyThrLeuIleProTyrArgCysAspLeuSerAsnGluGluAspIleLeuS 79
Db      307 GCTACCCCGGAGCTTTGATCCCCCTACAGATGTGACCTATCAATGAAGAGGACATCCTCT 366
QY      79 exMetPheSerAlaIleArgSerGlnHisSerGlyValAspIleCysIleAsnAsnAlaG 99
Db      367 CCATGTCTCAGCTATCCGTTCTCAGCAGCGGGTGTAGACATCTGCATCAACATGCTGTG 426
QY      99 lYLeuAlaArgProAsp-ThrLeuLeuSerGlySerThrSerGlyTyrLysAspMetPhe 118
Db      427 GCTGGCCCGGCTGACAAACCTGCTCTCAGCAGCAGCAGCAGTGGTGGAGGACATGTC 486
QY      119 Asn-VallAsnVal-LeuAlaLeuSerIleCysThrArgGluAlaTyrGlnSerMetLysG 138
Db      487 AATTGTGAACGTTGTGGCCCTCAGCATCTGCACACGGGAAGCCTACCACTCCATGAAGG 546
QY      138 luArgAsnValAspAspGlyHisIleIleAsnIleAsnSerMetSerGlyHisArgValL 158
Db      547 ACGGAATGTGACCATGGGCACATCAATTAACATCAATAGCATGTCTGGCCACCGAGTGT 606
QY      158 euProLeuSerValThrHisPheTyrSerAlaThrLysTyrAlaValThrAlaLeuThrG 178
Db      607 TACCCCTGTCTGTGACCCACTTCTATAGTGCACCAAGATGTCGGTCACTGACATGACAG 666
QY      178 luGlyLeuArgGlnGluLeuArgGluAlaGlnThrHisIleArgAlaThrCys-IleSer 197
Db      667 AGGGACTGAGGCAAGAGCTTCGGAGGGCCAGACCCAGATTCGAGCCACGTCGAATCTCT 726
QY      198 ProGlyValValGluThr-GlnPheAlaPheLysLeuHisAspLysAspProGluLysAl 217
Db      727 CCAGTGTGGGGAGACAAACAAATTCGCTTCAAACTCCACGACGAGCCCTTGAGAAGC- 785
QY      217 aAlaAlaThrTyrGluGlnMetLysCys-LeuLysProGluAspValAlaGluAlaValI 237
Db      786 -AGTGGCACTATGAGCAATCAAGGTGCTCAAAACCGAGGATGTGGCCGAGGCTGTAA 844
QY      237 leTyrValLeuSerThrProAlaHisIleGlnIleGlyAspIleGlnMetArgProThrG 257
Db      845 TCTACGTACTCAGGACC---CTGAACAATCAGAATGGAGACATCCAAATTAAGGCCAGGG 901
QY      257 luGln 258
Db      902 ACAGG 906

RESULT 8
BG967186
LOCUS
DEFINITION
602833867F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4988176 5',
mRNA sequence.
ACCESSION
BG967186
VERSION
BG967186.1 GI:14354823
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 914)
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.

```

cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11000 row: f column: 17
High quality sequence stop: 748.
Location/Qualifiers
1. .914
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4988176"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP Co24"
/note="Organ: colon; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

FEATURES
source

ORIGIN

Alignment Scores:
Pred. No.: 1,98e-110 Length: 914
Score: 1049.50 Matches: 223
Percent Similarity: 90.15% Conservative: 15
Best Local Similarity: 84.47% Mismatches: 20
Query Match: 78.50% Indels: 8
DB: 12 Gaps: 1

US-10-063-735-128 (1-260) x BG967186 (1-914)

QY 1 MetalArgProGlyMetGluArgTrpArgAspArgLeuAlaLeuValThrGlyAlaSer 20
Db 105 ATGACTAGAGCTGGCATGGAGCGGTGGCGCGACCGCTGGCACTGGTGACGGAGCCTCG 164
QY 21 GlyGlyIleGlyAla-AlaValAlaArgAlaLeuValGlnGlnGlyLeuLysValValG 40
Db 165 GGGGGCATCGGTGTGGCCGTGGCCCGGCGCATAGTCCAGCAGGAGCTGAAGGTGTGTGG 224
QY 40 YCysAlaArgThrValGlyAsnIleGluGluLeuAlaGluCysLys-SerAlaGlyT 60
Db 225 TTGTGCCCGCAGCCGTGGCAACATCGAGAGCTGGCTGCTGAATGTAGAGTGAGGCT 284
QY 60 YrProGlyThrLeuIleProTyrArgCysAspLeuSerAsnGluGluAspIleLeuSerM 80
Db 285 ACCCGGGACCTTGTATCCCTACAGATGTGACCTGTCAATGAGGAGGACATCTCTCCA 344
QY 80 etPheSerAlaIleArgSerGlnHisSerGlyValAspIleCysIleAsnAsnAlaGlyL 100
Db 345 TGTTCTCAGCTGTCCGATCCCGACAGTCAGTGGCTGATATCTGCATCAACATGCCGCA 404
QY 100 euAlaArgProAspThrLeuLeuSerGlySerThrSerGlyTyrLysAspMetPheAsnV 120
Db 405 TGGCCCGGCTTGACACCCCTGTCTCGGCAGACACCGAGCGGATGGAGGACATGTTCAATG 464
QY 120 alAsnValLeuAlaLeuSerIleCysThrArgGluAlaTyrGlnSer-MetLysGluArg 139
Db 465 TGAATGTGTGGCCCTCAGCATCTGCATCTGGGAGGCTTATCAGTCCACTCAGCGGCGG 524
QY 140 AsnValAspAspGlyHisIleIleAsnIleAsnSerMetSerGlyHisArgValLeuPro 159
Db 525 AACATAGACGAGCGGCACATCATTAACATCAACAGCATGTGTGGCCACCGAGTCCACCC 584
QY 160 LeuSerValThrHisPheTyrSerAlaThrLysTyrAlaValThrAlaLeuThrGluGly 179
Db 585 CAGTCTGTGATCCATTTCTATAGTGGGACTAAGTATGCCGTCACTGCACACTGACAGAGGA 644
QY 180 LeuArgGlnGluLeuArgGluAlaGlnThrHisIleArgAlaThrCysIleSerProGly 199
Db 645 CTCAGGCA-GAGCTTCTGGAGGCCACCACTATCCGGGC-ACGTGTATCTCTCCAGGC 702

QY 200 Val-ValGluThrGlnPheAlaPhelysleuHisAspPro---GluLysAlaAl 218
 Db ::|
 703 TTGCGTAGAGACACAGTTCGCTTCAATCCATGACAAGGACCCGGGGAACGACGCTG 762
 QY 218 aAlaThrTyTGluGlnMetLysCysLeuLysProGluAspValAlaGluAlaValile-T 238
 Db ::|
 763 CCCACCTATGATCACTTACCAGTGCTCAGACACGAGGAGCTGTCTGAGGCTGTCACT 822
 QY 238 yValLeuSerThrProAlaHisIleGlnIleGlyAspIleGlnMetArgProThrGluG 258
 Db ::|
 823 ACGTGTAGCAACACCCACATGTGAGGTTGGGACATCCAGATGAGGCCACAGACG 882
 QY 258 ln 258
 Db ::
 883 AG 884

RESULT 9
 BU539642
 LOCUS
 DEFINITION
 AGENCOURT_10224058 NIH_MGC_107 Homo sapiens cDNA clone
 IMAGE:6570429 5', mRNA sequence.

ACCESSION
 VERSION
 SOURCE
 ORGANISM
 Homo sapiens (human)

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

1 (bases 1 to 937)
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Prepared by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCM2759 row: e column: 21
 High quality sequence stop: 527.

FEATURES
 source
 1..937
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6570429"
 /tissue_type="adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_107"
 /note="Organ: breast; Vector: pOTB7; Site:1: EcoRI;
 Site:2: XhoI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

ORIGIN
 Alignment Scores:
 Pred. No.: 5,79e-109 Length: 937
 Score: 1037.00 Matches: 216
 Percent Similarity: 90.73% Conservative: 9
 Best Local Similarity: 87.10% Mismatches: 14
 Query Match: 77.56% Indels: 9
 DB: 13 Gaps: 3

US-10-063-735-128 (1-260) x BU539642 (1-937)
 QY 9 TrpArgAspArgLeuAlaLeuValThrGlyAlaSerGlyGlyIleGlyAlaValAla 28

Db
 2 TGGCGGACCGCTGGCGTGGTGGACGGGGGCTCGGGGGGCATCGCGCGCGCTGGCC 61
 QY
 29 ArgAlaLeuValGlnGlnGlyLeuLysValValGlyCysAlaArgThrValGlyAsnIle 48
 Db
 62 CGGGCCCTGGTTCAGCAGGAGCTGAAGTGTGGTGGCTGCGCCGCACTGTGGGCAACATC 121
 QY
 49 GluGluLeuAlaGluCysIlySerAlaGlyTyProGlyThrleuileProTyArg 68
 Db
 122 GAGGAGCTGGCTGCTGAATGTAAAGTGCAGGCTACCCGGGACTTTGATCCCTTACAGA 181
 QY
 69 CysAspLeuSerAsnGluAspIleLeuSerMetPheSerAlaIleArgSerGlnHis 88
 Db
 182 TGTGACCTATCAATGAAGAGGACATCTCTCCATGTTCTCAGTATCCGTTCTCAGCAC 241
 QY
 89 SerGlyValAspIleCysIleAsnAlaGlyLeuAlaArgProAspThrLeuLeuSer 108
 Db
 242 AGCGGTGTAGACATCTGCATCAACAATGTGGCTTGGCCCGGCTGACACCCCTGCTCTCA 301
 QY
 109 GlySerThrSerGlyTrpLysAspMetPheAsnValAsnValLeuAlaLeuSerIleCys 128
 Db
 302 GGCAGCACCGAGTGGTTGGAAGGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTGC 361
 QY
 129 ThrArgGluAlaTyxGlnSerMetLysGluArgAsnValAspAspGlyHisIleLeuAsn 148
 Db
 362 ACACGGGAGCCTACCAGTCCATGAAGAGCGGAATGTGACGATGGGCGACATCAATTAAC 421
 QY
 149 IleAsnSerMetSerGlyHisArgValLeuProLeuSerValThrHisPheTySerAla 168
 Db
 422 ATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGTCTGTGACCCACTTCTATAGTGCC 481
 QY
 169 ThrLysTyxAlaValThrAlaLeuThrGluGlyLeuArgGlnGlnLeuArgGluAlaGln 188
 Db
 482 ACCAAGTATGCCGCTCACTCGCTGACAGAGGAGCTGAGGCAAGAGCTTCGGGAGGCCAG 541
 QY
 189 ThrHisIleArgAlaThrCysIleSerProGlyValValGluThrGlnPhe-AlaPheTy 208
 Db
 542 ACCACATCCGAGCCACGTCATCTCTCCAGGTGGGTGGAGACCCCAATTCGCCCTTCCA 601
 QY
 208 sIleHisAsp-LysAspPro---GluLysAlaAlaAlaThrTyxGlu-GlnMetLysCys 226
 Db
 602 ACTCCACGAACAAAGGACCCCTCGAGAAAGGCTTCTGCCACACCTATGAACCAATGGAAGTG 661
 QY
 227 Leu---LysProGluAspValAlaGluAlaIleTyxValleu-----Ser 241
 Db
 662 GTCTCCAAACCCCTAAGAATGTCCCCCGAAGCTTGTATTCTACCGTCCCGCCCAAC 721
 QY
 242 ThrProAlaHisIleGln 247
 Db
 722 CCCCCGAAAAAATCCAG 739

RESULT 10
 BG030249
 LOCUS
 DEFINITION
 602297547F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:4392011 5',
 mRNA sequence.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 932)
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCTD/DTF
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLMW10083 row: n column: 12
High quality sequence stop: 601.
Location/Qualifiers
1. .932
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4392011"
/issue_type="mammary adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 87"
/note="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dr primed. Average insert size 1.383 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

FEATURES

source

ORIGIN

Alignment Scores:
Pred. No.: 1.41e-107 Length: 932
Score: 1025.00 Matches: 206
Percent Similarity: 97.18% Conservative: 1
Best Local Similarity: 96.71% Mismatches: 5
Query Match: 76.66% Indels: 2
DB: 10 Gaps: 0

US-10-063-735-128 (1-260) x BG030249 (1-932)

QY 49 GluGluLeuAla-AlaGluCysLysSerAlaGlyTyrProGlyThrLeuLeuProTyrAr 68
DB 1 GAGGAGCTGGCTGCTGAATGTAAGATGCAGCTACCCCGGAGCTTTGATCCCTACAG 60
QY 68 gCysAspLeuSerAsnGluAspIleLeuSerMetPheSerAlaIleArgSerGlnHi 88
DB 61 ATGTGACCTATCAATGAAGAGGACATCTCTCCATGTTCTCAGCTATCCGTTCTCAGCA 120
QY 88 sSerGlyValAspIleCysIleAsnAsnAlaGlyLeuAlaArgProAspThrLeuLeuSe 108
DB 121 CAGCGGTGTAGACATCTGCATCAACAATGTGGCTTGGCCGCGCTGACACCCCTGCTCTC 180
QY 108 rGlySerThrSerGlyTyrLysAspMetPheAsnValAsnValLeuAlaLeuSerIleCy 128
DB 131 AGGCAGCACCGTGGTTGAAGAGCATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTG 240
QY 128 sThrArgGluAlaTyrGlnSerMetLysGluArgAsnValAspAspGlyHisIleIleAs 148
DB 241 CACACGGGAAGCTTACCAGTCCATGAAGAGCGGAATGTGGACGATGGGCACATCATTA 300
QY 148 nIleAsnSerMetSerGlyHisArgValLeuProLeuSerValThrHisPheTyrSerAl 168
DB 301 CATCAATAGCATGTCTGGCCACCGAGTGTATCCCTGTCTGTGACCCCACTTCTATAGTC 360
QY 168 aThrLysTyrAlaValThrAlaLeuThrGluGlyLeuArgGlnGlnLeuArgGluAlaGl 188
DB 361 CACCAAGTATGCCGTACCTGCGCTGACAGAGGATGTGAGCAAGAGCTTGGGAGGCCCA 420
QY 188 nThrHisIleArgAlaThrCysIleSerProGlyValValGluThrGlnPheAlaPheLy 208
DB 421 GACCCACATCCGAGCCACGTGCATCTCTCCAGTGTGGTGGAGACACAATTCGCTCTCA 480
QY 208 sLeuHisAspLysAspProGluLysAlaAlaThrTyrGluGlnMetLysCysLeuLy 228
DB 481 ACTCCACGACAAAGGACCTCTGAGAGGACAGTCCACCTATGAGCAAAATGAAGTGTCTCA 540
QY 228 sProGluAspValAlaGluAlaValIleTyrValLeuSerThrProAlaHisIleGlnI 248
DB 541 ACCCGAGGATGTGGC-GAGGCTGTATTCTTACGTCTCTCAGCACCCCGCACATCCAGAT 599
QY 248 eGlyAspIleGlnMetArgProThrGluGlnValThr 260
||||| :||| |||||

Db 600 TGGAGACATCAGATGAGGGCCACGAGCAGGTGACC 636

RESULT 11

BE796469

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1. .1063
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3943948"
/issue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_7"
/note="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dr priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

Alignment Scores:
Pred. No.: 4.61e-104 Length: 1063
Score: 995.50 Matches: 225
Percent Similarity: 88.30% Conservative: 9
Best Local Similarity: 84.91% Mismatches: 26
Query Match: 74.46% Indels: 9
DB: 10 Gaps: 1

US-10-063-735-128 (1-260) x BE796469 (1-1063)

QY 1 MetAlaArgProGlyMetGluArgTyrArg-AspArgLeuAlaLeuValThrGlyValaSe 20
DB 135 ATGGCCAGGCCCGCATGAGCGGTGGCGCTGACCGCTGGCGCTGGTGCAGCGGGGCCCTC 194
QY 20 rGlyGlyIleGlyAlaAlaValAlaArgAlaLeuValGlnGlnGlyLeuLysValValGl 40
DB 195 GGGGGGCATCGCGCGCGCGTGGCGCCCGCTGCTGCAGCAGGAGCTGAAGGTGGTGGG 254
QY 40 yCysAlaArgThrValGlyAsnIleGluGluLeuAlaAlaGluCysLysSerAlaGlyTy 60
DB 255 CTGGCCCCGCACTGTGGGCAACATCGAGAGCTGCTGCTGAATGTAAGAGTGCAGGCTA 314
QY 60 rProGlyThrLeuIleProTyrArgCysAspLeuSerAsnGluGluAspIleLeuSerMe 80
DB 315 CCCCAGGACTTTGATCCCTTACAGTGTGACCTATCAATGAAGAGGACATCTCTCCAT 374


```

QY      80 tPheSerAlaIleArgSerGlnHisSerGlyValAspIleCysIleAsnAsnAlaGlyLe 100
Db      375 GTTCTCAGCTATCCGTTCTCAGCAGCGGTTAGACATCTGCATCAACAATGCTGGCTT 434

QY      100 uAlaArgProAspThrLeuLeuSerGlySerThrSerGlyTyrLysAspMetPheAsnVa 120
Db      435 GCGCCGGCTGACACCTCTCTCAGGAGCCAGCTGTTGGAAGGACATGTTCAATGT 494

QY      120 lAsnValLeuAlaLeuSerIleCysThrArgGluAlaTyrGlnSerMetLysGluArgAs 140
Db      495 GAAACGTGCTGGCCCTCAGCATCTGCACACGGAAGCCTACCAAGTCCATGAGGAGCGGAA 554

QY      140 nValAspAspGlyHisIleIleAsnIleAsnSer-MetSerGlyHisArgValLeuProL 160
Db      555 TTGTGACGATGGGCACATCAATTAACATCAATAGCATTTGTCGGCACCGAGTGTACCGC 614

QY      160 euSerValThrHisPheThrAlaThrLysThrAlaValThrAlaLeuThrGluGly- 179
Db      615 TGTCTGTGACCAA-TTCTATAGTGCCAC-AAGTATGCCGTCACTGGCTGACAGGGAC 672

QY      180 LeuArgGlnGluLeuArgGluAlaGlnThrHisIleArgAlaThrCysIle-SerProGl 199
Db      673 TTGAGCGAAGAGCTTCGGAGGCGCAAGACCCACATCCGAGCCACGTCGATTTCTCCAGG 732

QY      199 vValValGluThrGlnPheAlaPheLysLeuHisAspLysAspProGluLysAlaAlaL 219
Db      733 TTGTGGGAGACACAATTCCGCTTCAAAAGTCCA-GACAAGGACCCCTGAGAAGGGCAGTGG 791

QY      219 aThrYrGluGlnMetLysCysLeuLysProGluAspValAlaGluAlaVal---IleTy 238
Db      792 CAACATATAAGCAATTGAGTGG-TTAAACCCGAGATTTTGCCGAGGATGGAATTAAAG 850

QY      238 rValLeuSerThrProAlaHisIleGlnIleGlyAspIleGlnMetArgProThrGluGl 258
Db      851 GGTTCATCAGAAACCCCGCAAGATTCGGAATGCGGAATTCACAATAATGCCACGAGAG 910

QY      258 nValThr 260
Db      911 CCGGACC 917

RESULT 12
BX671560
LOCUS      BX671560          612 bp      mRNA      linear      EST 24-OCT-2003
DEFINITION BX671560 Sus Scrofa library (scac) Sus scrofa cDNA clone
            scac0032i.1.i.13 5prim, mRNA sequence.
ACCESSION  BX671560
VERSION     BX671560.1
KEYWORDS   EST.
SOURCE     Sus scrofa (pig)
ORGANISM   Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE  1 (bases 1 to 612)
AUTHORS   Bonnet,A., Tosser-Klopp,G., Benne,F., Cabau,C., Villegier,S.,
            Soares,M., Bonaldo,P. and Hately,F.
TITLE     A Pig Normalised Multi-Tissue cDNA Library
JOURNAL   Unpublished (2003)
COMMENT   Contact: Tosser-Klopp G
            Genetique Animale
            Institut National de la Recherche Agronomique
            Chemin de Borde-Rouge - Auzeville BP27, 31326 Castanet-Tolosan
            cedex, FRANCE
            Tel: 33 (0) 5.61.28.51.14
            Fax: 33 (0) 5.61.28.53.08
            Email: tosser@toulouse.inra.fr
            Clone distribution: AGENAE Resource centre, Francois PIUMI,
            Francois.Piumi@jouy.inra.fr, INRA, CEA Radiobiologie et Etude du
            genome (LREG), Domaine de Vilvert, 78352, Jouy-en-Josas cedex,
            FRANCE, +33 (0) 1.34.65.28.02, +33 (0) 1.34.65.22.73.
            Sequence cleaned of vector, adaptor and repetitions. Contact us
            at signenasupport@jouy.inra.fr to obtain the chromatogram of this
            sequence.

```

```

Plate: 0032 row: 1 column: 13.
FEATURES
    source
        1..612
            /location/Qualifiers
            /organism="Sus scrofa"
            /mol_type="mRNA"
            /db_xref="taxon:9823"
            /clone="scac0032i.1.i.13"
            /tissue_type="mixed"
            /clone_lib="Sus Scrofa library (scac)"
            /note="Vector: pT73D-pac vector; tissues: adipose tissue,
            brain, kidney, liver, muscle, ovary, testis, heart,
            hypothalamus, pancreas, skin, spleen, thymus, placenta,
            pituitary gland, seminal vesicle, small intestine,
            uterus, adrenals, bulbo uretral gland, cerebral trunk,
            epididymis, female gonad, gall-bladder, hippocampus,
            large intestine, male gonad, melanocytes, stomach, udder"

```

ORIGIN

```

Alignment Scores:
Pred. No.:      4,91e-103      Length:      612
Score:          983.00         Matches:    193
Percent Similarity: 97.55%      Conservative: 6
Best Local Similarity: 94.61%    Mismatches:  4
Query Match:    73.52%         Indels:      1
DB:             13             Gaps:        0

US-10-063-735-128 (1-260) x BX671560 (1-612)

QY      12 ArgLeuAlaLeuValThrGlyAlaSerGlyGlyIleGlyAlaAlaValAlaArgAlaLeu 31
Db      1  CGACTGGCACATGTTGCACAGGACCTCGGGGGGCATCGGGCGCGCTGCGCCCGGGCCCTG 60

QY      32 ValGlnGlnGlyLeuLysValValGlyCysAlaArgThrValGlyAsnIleGluGluLeu 51
Db      61 GTCCAGCAGCGACTCAAGGTGTTGGTGTGTGCGCCGACCGTGGGCAACATCGAGGAGCTG 120

QY      52 AlaAlaGluCysLysSerAlaGlyTyrProGlyThrLeuIleProTyrArgCysAspLeu 71
Db      121 GCGGCGGAGTGTAAAGTGCAGGCTACCCCGGAGCTTTGATCCCTTACAAATGTGACCTG 180

QY      72 SerAsnGluGluAspIleLeuSerMetPheSerAlaIleArgSerGlnHisSerGlyVal 91
Db      181 TCCAAATGAGGAGGACATCTCTCCATGTTCTCGGCCATCCGCTCTCAACACAGCGGTGTA 240

QY      92 AspIleCysIleAsnAsnAlaGlyLeuAlaArgProAspThrLeuLeuSerGlySerThr 111
Db      241 GACATCTGCATCAACAACGCGGGCTTGGCCCGCCGACAGCTGCTCTCAGGCGAGCACC 300

QY      112 SerGlyTyrLysAspMetPheAsnValLeuAlaLeuSerIleCysThrArgGlu 131
Db      301 AGCGGTTGGAAGGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACCCGGGAA 360

QY      132 AlaTyrGlnSerMetLysGluArgAsnValAspAspGlyHisIleIleAsnIleAsnSer 151
Db      361 GCCTACCAAGTCCATGAGGAGCGCAAGGTGGATGATGGGCACATCATTAACATCAACAGC 420

QY      152 MetSerGlyHisArgValLeuProLeuSerValThrHisPheTyrSerAlaThrLysTyr 171
Db      421 ATGCTGCGCCACCGAGTGTGATACCCCGAGTCTGTGGGCCCATTTCTATAGTGTACCAAGTAT 480

QY      172 AlaValThrAlaLeuThrGluGlyLeuArgGlnGluLeuArgGluAlaGlnThrHisIle 191
Db      481 GCGGTCATCTGCGTCACAGAGGGCTTGAGGAGAGAGCTTCGGGAGGCCAGACCCACATC 540

QY      192 ArgAlaThrCysIleSerProGlyValValGluThrGlnPheAlaPheLysLeu-HisAs 211
Db      541 CGAGCAGCGTGCATCTCTCCAGGAGTCTGTGGAGACACAGATTCGCTTCAAAATCCCATCA 600

QY      211 playspPro 214
Db      601 CCAGAACCCCC 610

```

RESULT 13

```

CK019037
LOCUS       CK019037               807 bp    mRNA    linear    EST 26-NOV-2003
DEFINITION  AGENCOURT_16543761 NIH_ZGC_10 Danio rerio cDNA clone IMAGE:7041546
5', mRNA sequence.
ACCESSION   CK019037
VERSION     CK019037.1   GI:38544961
KEYWORDS    EST.
SOURCE      Danio rerio (zebrafish)
ORGANISM    Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1 (bases 1 to 807)
            NIH-MGC http://mgi.nci.nih.gov/
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Len Zon, Harvard
            cDNA Library Preparation: Open Biosystems
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM14798 row: c column: 16
            High quality sequence stop: 686.

FEATURES             source
    1..807
        /organism="Danio rerio"
        /mol_type="mRNA"
        /db_xref="taxon:7955"
        /clone="IMAGE:7041546"
        /tissue_type="whole body"
        /lab_host="DH10B"
        /clone_lib="NIH_ZGC_10"
        /note="Vector: pExpress1; Site_1: NotI; Site_2: EcoRV;
            Bulk tissue was collected from a whole adult individual
            from the Tuebingen strain. 1st strand cDNA was primed with
            a Not I - oligo(dT) primer, double-stranded cDNA was
            cloned into the Not I and EcoRV sites of pExpress-1.
            Library was size-selected for >1 kb fragments. A
            normalized version of this library is also available
            (NIH_ZGC_7). Library was constructed by Open Biosystems
            (Huntsville, AL)."
```

Alignment Scores:

Pred. No.:	2.81e-100	Length:	807
Score:	961.00	Matches:	178
Percent Similarity:	85.49%	Conservative:	40
Best Local Similarity:	69.80%	Mismatches:	37
Query Match:	71.88%	Indels:	0
DB:	14	Gaps:	0

US-10-063-735-128 (1-260) x CK019037 (1-807)

```

QY      6 MetGluArgTgTpArgAspArgLeuAlaLeuValThrGlyAlaSerGlyGlyIleGlyAla 25
DB      12 ATGGATCGCTGGAAGGACAGAGTCTCTGTGCTGGAGCTTCAGTGAATAGAGCT 71
QY      26 AlaValAlaArgAlaLeuValGlnGlnGlyLeuLysValValGlyCysAlaArgThrVal 45
DB      72 GCAATCGCAAAAGCTCTTGTCCAGCATGCATGAGAGGTGGTGGATGTGCCAGAAATGTG 131
QY      46 GlyAsnIleGluGluLeuAlaGluCysLysSerAlaGlyTyrProGlyThrLeuIle 65
DB      132 GAGCAAAATAGAGAAATCGCGCTGAATGTGTCACTGCGGACTCAGCGGTCTTCTGTT 191
QY      66 ProTyrArgCysAspLeuSerAsnGluGluAspIleLeuSerMetPheSerAlaIleArg 85

192 CCATATAATGTGATCTTTCAGTAGAGGATGAAGTTTGTCCATGTTCTCTCGGATTAG 251
86 SerGlnHisSerGlyValAspIleCysIleAsnAsnAlaGlyLeuAlaArgProAspThr 105
252 GCTCAACATAAGGGTGTGTGATGATTAATAATGCTGGTTAGCTCTGCCAGAGCCT 311
106 LeuLeuSerGlySerThrSerGlyTropLysAspMetPheAsnValAsnValLeuAlaLeu 125
312 CTTGTTGAACGGCAAGCCAGCGGCTGGAGGACTATGATGAACGTAAGTAAATGTAATGGCCTG 371
126 SerIleCysThrArgGluAlaTyrGlnSerMetLysGluArgAsnValAspAspGlyHis 145
372 GCAGTGTGACCCCGTGAGGCTTACCAGTCCATGAAGAAGAAAGAAATATTGATGATGGCCAT 431
146 IleIleAsnIleAsnSerMetSerGlyHisArgValLeuProLeuSerValThrHisPhe 165
432 ATCATTAATATTAAACAGTATGTCTGGCATCGGTTGTAAACAGTCGCTATACACATTC 491
166 TyrSerAlaThrLysTyrAlaValThrAlaLeuThrGluGlyLeuArgGlnLeuArg 185
492 TACACCGCTACTAAATACGCAGTGACTCTCTCACCGAAGGTTTGAGGCAAGAGTTACGA 551
186 GluAlaGlnThrHisIleArgAlaThrCysIleSerProGlyValValGlnThrGlnPhe 205
552 GAGGCCAAACCCCATACGTGCGCACAGTATATCCCTGGTTTAGTGGAGACAGAAATT 611
206 AlaPheLysLeuHisAspLysAspProGlyLysAlaAlaAlaThrTyrGluGlnMetLys 225
612 GCCTACAGACTCTTTAGTGAACCAACAGACAGGCTTCTGCCACCCTACAAAGTATAAG 671
226 CysLeuLysProGluAspValAlaGluAlaValIleTyrValLeuSerThrProAlaHis 245
672 TGCTTGCCAGCCAGATGATCTAGCANATGAGTGGTTTATGTCCTAAGCGCTCTCCCTCA 731
246 IleGlnIleGlyAspIleGlnMetArgProThrGluGlnValThr 260
732 TGTTCAATTGGTGACATTCAGATGAGACCTGTGGACAGCTGACA 776

RESULT 14
CF221812
LOCUS     CF221812               769 bp    mRNA    linear    EST 04-AUG-2003
DEFINITION  AGENCOURT_14994658 NIHCD_XGC_Emb5 Silurana tropicalis cDNA clone
IMAGE:6986808 5', mRNA sequence.
ACCESSION   CF221812
VERSION     CF221812.1   GI:33422520
KEYWORDS    EST.
SOURCE      Silurana tropicalis (western clawed frog)
ORGANISM    Silurana tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
            Xenopodinae; Silurana.
REFERENCE   1 (bases 1 to 769)
            NIH-MGC http://mgi.nci.nih.gov/
AUTHORS     National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE       Unpublished (1999)
JOURNAL     Contact: Daniela S. Gerhard, Ph.D.
COMMENT     Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Robert M. Grainger
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM14656 row: j column: 23
            High quality sequence start: 3
            High quality sequence stop: 711.

FEATURES             source
    1..769
```

```
/organism="Silurana tropicalis"  
/mol_type="mRNA"  
/db_xref="taxon:8364"  
/clone="IMAGE:6986808"  
/tissue_type="gastrula"  
/dev_stage="embryo, stages 10-13"  
/lab_host="DH10B (phage-resistant)"  
/clone_lib="NICHD_XGC_Emb5"  
/notes="Vector: pCMV-SF0R6.1; Site 1: NotI; Site 2: EcoRV;  
Cloned unidirectionally. Primer: Oligo dt. Average insert  
size 2.0 kb. Constructed by Invitrogen. Note: This is a  
Xenopus Gene Collection (XGC) library."
```

ORIGIN

```
Alignment Scores:  
Pred. No.: 9,81e-100 Length: 769  
Score: 956.00 Matches: 182  
Percent Similarity: 89.04% Conservative: 21  
Best Local Similarity: 79.82% Mismatches: 25  
Query Match: 71.50% Indels: 0  
DB: 14 Gaps: 0
```

US-10-063-735-128 (1-260) x CP221812 (1-769)

```
QY 6 MetGluArgTrrArgAspArgLeuAlaLeuValThrGlyAlaSerGlyGlyIleGlyAla 25  
DB 82 ATGGAGCGTGAAGGCGAGGTGGACCTTTGACCGGGCTCGGTGGCATCGAGCC 141  
QY 26 AlaValAlaArgAlaLeuValGlnGlnGlyLeuLysValValGlyCysAlaArgThrVal 45  
DB 142 GCGGTGCGCGGTGCTGTTCAGCATGTCATGAAGTGGTGGCTGCCAGGCGT 201  
QY 46 GlyAsnIleGluCluLeuAlaGluCysLysSerAlaGlyTyrProGlyThrIleu 65  
DB 202 GATAAGATTTGAAACTGCTGTAATGTGAGGTGGTACCAGGCACCTTATT 261  
QY 66 ProTyrArgCysAspLeuSerAsnGluAlaGlyLeuSerMetPheSerAlaIleArg 85  
DB 262 CTTTAAATGTGACCTGCTCAATGAAGAGGAGATTCCTCATGTTTCAGCAATAAG 321  
QY 86 SerGlnHisSerGlyValAspIleCysIleAsnAsnAlaGlyLeuAlaArgProAspThr 105  
DB 322 ACTTTGCATCAGGGGTGCGATGTATGATCAACATGCAGGTGGCCGACCGAGCCT 381  
QY 106 LeuLeuSerGlySerThrSerGlyTrrLysAspMetPheAsnValAsnValLeuAlaLeu 125  
DB 382 TTGCTGAGTGGCAAAACAGAGGGATGGAGAACAAATGATTGATTAAATGTTCTTGCACTC 441  
QY 126 SerIleCysThrArgGluAlaTrrGlnSerMetLysGluArgAsnValAspAspGlyHis 145  
DB 442 AGTATCTGCACAGAGAGCGCTTACCATGTCATGAAGGAAGGAATATCGATGCGCAT 501  
QY 146 IleIleAsnIleAsnSerMetSerGlyHisArgValLeuProLeuSerValThrHisPhe 165  
DB 502 ATCATAAACATCAACAGCATGATGGCCATAGAGTTCTTCTTCTACACTTATGCACTTT 561  
QY 166 TyrSerAlaThrIlystyrrAlaValThrAlaLeuThrGluGlyLeuArgGlnGluLeuArg 185  
DB 562 TATTTCAGCTACTTAAGTATGCTGTAACTGCGCTGACAGAGGGCTCAGGAAGAGTTCAGA 621  
QY 186 GluAlaGlnThrHisIleArgAlaThrCysIleSerProGlyValValGluThrGlnPhe 205  
DB 622 GAAGAAAGAGTCACATCCGAGCAACAGATATATCCGAGGCGCTTGTGGAACCTGATTT 681  
QY 206 AlaPheLysLeuHisAspLysAspProGluLysAlaAlaAlaThrTrrGluGlnMetLys 225  
DB 682 GCATTTAAACTCTTGTATGATCCGGAANGCTGCTGNCACATATGAAAGTATAAG 741  
QY 226 CysLeuLysProGluAspValAla 233  
DB 742 TGCTGGAAGCTGAAGACATTTGCT 765
```

RESULT 15

```
BE617298  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT
```

```
BE617298 923 bp mRNA linear EST 20-OCT-2000  
60141804F1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3846168 5',  
mRNA sequence.  
BE617298  
BE617298.1 GI:9888236  
EST.  
Homo sapiens (human)  
Homo sapiens  
Mammalia; Euthera; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 923)  
NIH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM9558 row: 9 column: 01  
High quality sequence stop: 698.
```

FEATURES
source

```
1..923  
/organism="Homo sapiens"  
/mol_type="mRNA"  
/db_xref="taxon:9606"  
/clone="IMAGE:3846168"  
/tissue_type="adenocarcinoma"  
/lab_host="DH10B (phage-resistant)"  
/clone_lib="NIH_MGC_65"  
/note="Organ: colon; Vector: pCMV-SF0R6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.8 kb. Library constructed by Life  
Technologies."
```

ORIGIN

```
Alignment Scores:  
Pred. No.: 1,76e-99 Length: 923  
Score: 955.00 Matches: 186  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 71.43% Indels: 0  
DB: 10 Gaps: 0  
US-10-063-735-128 (1-260) x BE617298 (1-923)  
QY 75 GluAspIleLeuSerMetPheSerAlaIleArgSerGlnHisSerGlyValAspIleCys 94  
DB 3 GAGGACATCTCTCCATGTTCTCAGCATTCCTTCAGCAGCAGCGGTGTAGACATCTGC 62  
QY 95 IleAsnAsnAlaGlyLeuAlaArgProAspThrLeuSerGlySerThrSerGlyTrr 114  
DB 63 ATCAACAATGCTGGCTTGGCCGGCTGACACCTCTCTCAGGAGCAGCAGTGTGG 122  
QY 115 LysAspMetPheAsnValAsnValLeuAlaLeuSerIleCysThrArgGluAlaTrrGln 134  
DB 123 AAGGACATGTTCAATGTGAACGTGTGCCCTCAGCATCTGCACACGGAGACCTACCAG 182  
QY 135 SerMetLysGluArgAsnValAspAspGlyHisIleIleAsnIleAsnSerMetSerGly 154  
DB 183 TCCATGAAGGAGCGGAATGTGACATGGGCATCATTAACATCAATAGCATGTCTGGC 242  
QY 155 HisArgValLeuProLeuSerValThrHisPheTyrSerAlaThrLysTyrAlaValThr 174  
DB 243 CACCGAGTGTACCCCTGCTGTGACCCACTTCTATAGTGCACCAAGTATGCCGTCACT 302  
QY 175 AlaLeuThrGluGlyLeuArgGlnGluLeuArgGluAlaGlnThrHisIleArgAlaThr 194
```

Db 303 GCCTGACAGGAGCTGAGGCAAGAGCTTCGGGAGGCCCGAGACCACATCCGAGCCACG 362

QY 195 CysIleSerProGlyValValGluThrGlnPheAlaPheIleuHisAspLysAspPro 214

Db 363 TGCACTCTCCAGGTGGTGGAGACAAATTCGCTTCAAACTCCACGACAGGACCT 422

QY 215 GluLysAlaAlaAlaThrTyrGluGlnMetLysCysLeuLysProGluAspValAlaGlu 234

Db 423 GAGAGGACAGCTGCCACCTATGAGCAATCAAGTGTCTCAACCCGAGGATGTGGCCGAG 482

QY 235 AlaValIleTyrValLeuSerThrProAlaHisIleGlnIleGlyAspIleGlnMetArg 254

Db 483 GCTGTATCTACCTCTCAGCAGCCCGCAGCACATCCAGATTGGAGACATCCAGATGAG 542

QY 255 ProThrGluGlnValThr 260

Db 543 CCCACGAGCAGGTGACC 560

RESULT 16

LOCUS BG282361

DEFINITION BG2402863F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:4545231 5',

ACCESSION BG282361

VERSION BG282361.1 GI:13031288

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 852)

TITLE NIH-MGC http://mgi.nci.nih.gov/.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC/DCTD/DTP

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: LLCM1229 row: n column: 16

High quality sequence stop: 851.

FEATURES

source

1..852

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4545231"

/tissue_type="melanotic melanoma"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_20"

/note="organ: skin; Vector: pOTB7; Site: 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Alignment Scores:

Pred. No.: 8,67e-99 Length: 852

Score: 948.50 Matches: 206

Percent Similarity: 81.13% Conservativeness: 9

Best local Similarity: 77.74% Mismatches: 16

Query Match: 70.94% Indels: 34

DB: 12 Gaps: 3

US-10-063-735-128 (1-260) x BG282361 (1-852)

QY 1 MetAlaAtpProGlyMetGluArgTTPArgAspArgLeuAlaLeuValThrGlyAlaSer 20

Db 123 ATGGCCAGGCCCGGATGGAGCGGTGGCCGACCCGGCTGGTGGTGGAGCGGGGCTCG 182

QY 21 GlyGlyIleGlyAlaAlaValAlaArgAlaLeuValGlnGlnGlyLeuLysValValGly 40

Db 183 GGGGGCATCGCGCGCGCGCTGGCCGGGCTTGGTTCAGCAGGAGACTGAAGGTGGTGGGC 242

QY 41 CysAlaArgThrValGlyAsnIleGluGluLeuAlaAlaGluCysLysSerAlaGlyTyr 60

Db 243 TGGCCCGCACATGTGGCAACATCGAGGAGCTGGCTGTAATGTAAAGAGTGCAGGCTAC 302

QY 61 ProGlyThrLeuIleProTyrArgCysAspLeuSerAsnGlu-GluAspIleLeuSerMe 80

Db 303 CCCGGACTTTGATCCCTTACAGATGTGACCTATCAATGACCGAGGACATCTCTCCAT 362

QY 80 tPheSerAlaIleArgSerGlnHisSerGlyValAspIleCysIleAsnAsnAlaGlyLe 100

Db 363 GTTCTCAGCTATCCGTTCTCAGCAGCGGTGTAGACATCTGCATCAACATGTGGCTT 422

QY 100 uAlaArgProAspThrLeuLeuSerGlySerThrSerGlyTyrLysAspMetPheAsnVa 120

Db 423 GGCCCGGCTGACACCTGATCTCAGGCGAGCACCAGTGGTGGAGGACATGTTCAATCT 482

QY 120 lAsnVal-LeuAlaLeuSerIleCysThrArgGluAlaTyrGlnSerMetLysGluArg 140

Db 483 GAACGTGCTTGGCCCTCAGCATCTCAGCAGCGGAAGCCTTACCAGTCCATGAAGGAGCGGA 542

QY 140 snValAspAspGlyHisIleIleAsnIleAsnSerMetSerGlyHisArgValLeuProL 160

Db 543 CTGTGAGCAGATGGGCACATCATTAACATCAATAGCATGTCTGGACACCGAGTGTACCC 602

QY 160 euSerValThrHisPheTyrSerAlaThrLysTyrAlaValThrAlaLeuThrGluGlyL 180

Db 603 TGTCTGTGAACCATCTTATAGTGCACAAAGTATGCCGTCTACTGCGCTGACAGAGGAC 662

QY 180 euArgGlnGluLeuArgGluAlaGlnThr-HisIleArgAlaThrCysIle-SerProGl 199

Db 663 TGAGCAAGAGCTTCGGGAGGCCAGCCCAAAATCCGAGCCCAAGTGTCTCCAAACCCGG 722

QY 199 yValValGluThrGlnPheAlaPheLysLeuHisAspLysAspProGluLysAlaAla 219

Db 723 AGGAGTGTGGCCGAG---GCTGTAACTACGATCTCAGGAACCC-----GCCAA 770

QY 219 aThrTyrGluGlnMetLysCysLeuLysProGluAspValAlaGluAlaValIleTyrVa 239

Db 771 CACATTC----- 777

QY 239 lLeuSerThrProAlaHisIleGlnIleGlyAspIleGlnMetArgProThrGlu-GlnV 259

Db 778 -----CATATTGGACACATCCAGATGAGGGCCCGCAGGACGAGG 815

QY 259 alThr 260

Db 816 TGACC 820

RESULT 17

LOCUS CA474188

DEFINITION IMAGE:6795203 5', mRNA sequence.

ACCESSION CA474188

VERSION CA474188.1 GI:24930540

KEYWORDS EST.

SOURCE Danio rerio (zebrafish)

ORGANISM Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

REFERENCE 1 (bases 1 to 846)

AUTHORS NIH-MGC http://mgi.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: Leonard I. Zon, M.D.
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM14304 row: 0 column: 10
High quality sequence stop: 606.

FEATURES
Location/Qualifiers
1..846
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:6795203"
/lab_host="DH10B (T1-resistant)"
/clone_lib="NCI CGAP_ZKId1"
/note="Organ: kidney; Vector: pCMV-SPORT6.1; Site 1: ScovR; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dr. Average insert size 1.8 Kb. Constructed by J. Wang (Research Genetics, Invitrogen Corp) from tissue donated by L. Zon (Harvard University). Note: this is a NCI CGAP Library."

ORIGIN
Alignment Scores:
Pred. No.: 1.67e-98 Length: 846
Score: 946.00 Matches: 181
Percent Similarity: 86.05% Conservative: 41
Best Local Similarity: 70.16% Mismatches: 33
Query Match: 70.76% Indels: 3
DB: 14 Gaps: 0

US-10-063-735-128 (1-260) x CA474188 (1-846)

QY 6 MetGluArgTTPArgAspArgLeuAlaLeuValThrGlyAlaSerGlyGlyAla 25
DB 16 ATGGATCGCTGGAAGCGCAGAGTGTCTTGTCACCTGGAGCTTCAGTAGAATAGAGCT 75
QY 26 AlaValAlaArgAlaLeuValGlnGlnGlyLeuLysValValGlyCysAlaArgThrVal 45
DB 76 GCATTCGCAAAAGCTCTTCTCCAGCATGCGATGCAAGTGTGCGATGTCGAGATG 135
QY 46 GlyAsnIleGluGluLeuAlaGluCysLysSerAlaGlyTyrProGlyThrLeuIle 65
DB 136 GAGCAAAATAGAAAACCTGGCGCTGAATGTGTGAGTGGCGGACTCAGCGGTGCTCTGTTT 195
QY 66 ProTyrArgCysAspLeuSerAsnGluGluAspIleLeuSerMetPheSerAlaIleArg 85
DB 196 CCATATAAATGTGATCTTTCAGTAGAGGATGAAGTTTGTCCAGTGTCTCGGATTAAG 255
QY 86 SerGlnHisSerGlyValAspIleCysIleAsnAsnAlaGlyLeuAlaArgProAspThr 105
DB 256 GCTCAACATAAGGGTGTGTATGTTTCATTAATATGCTGTTTAGCTCTGCCAGAGCT 315
QY 106 LeuLeuSerGlySerThrSerGlyThrLysAspMetPheAsnValSerValLeuAlaLeu 125
DB 316 CTGTTGAACGGCAAAACCAAGCGCTGAGGACTATGATGAACGTGAATGTAATTCGCTG 375
QY 126 SerIleCysThrArgGluAlaTyrGlnSerMetLysGluArgAsnValAspGlyHis 145
DB 376 GCAGTGTGCCCGTGGGCTTACCAGTCCATGAAAGAAAGAAATATTGATGATGCCAT 435
QY 146 IleIleAsnIleAsnSerMetSerGlyHisArgValLeuProLeuSerValThrHisPhe 165
DB 436 ATCATTAATATTAAACAGTATGTCTGGGCATCGGGTTGTAACAGAGTCCATACACTTC 495
QY 166 TyrSerAlaThrLysTyrAlaValThrAlaLeuThrGluGlyLeuArgGlnGluLeuArg 195
DB 496 TACACCGCTACTAAATACGAGTGTGCTCTCACCGAAGGTTTGAGCGAAGAGTTACGA 555

QY 186 GluAlaGlnThrHisIleArgAlaThrCysIleSer-ProGlyValValGluThrGlnPh 205
DB 556 GAGGCCAAACCCACATCGTCCACAGTATATCCNCTGGTTAGTGGACAGAAAT 615
QY 205 eAlaPheLysLeuHisAspLysAspProGluLysAlaAlaAlaThrTyrGluGlnMetly 225
DB 616 TGCCTACAGACTCTTTAGTGAACCAAGACAAGGCTTCTGCCACTACAAAAGTATAA 675
QY 225 sCysLeuLysProGluAspValAlaGlu-AlaValIleTyrValLeuSerThrProAlaH 245
DB 676 GTGCCTGACGCGAGATGATCTAGCAATGGCAGTGGTTATGCTCCTAGCGCTCTCCTC 735
QY 245 IleGlnIleGlyAspIleGlnMetArg-ProThrGluGlnValThr 260
DB 736 ATGTTCAAAATGGTGACATTCAGATGAGACCTGTGGACACGCTGACA 783

RESULT 18
CA964199
LOCUS
DEFINITION CA964199 947 bp mRNA linear EST 03-JAN-2003
10d20 5', mRNA sequence.
ACCESSION CA964199.1 GI:27490756
VERSION CA964199.1
KEYWORDS EST.
SOURCE Cyprinus carpio (common carp)
ORGANISM Cyprinus carpio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Cyprinus.
1 (bases 1 to 947)
Gracey,A.Y., Fraser,E., Li,W. and Cossins,A.R.
Microarray and EST analysis of the carp (Cyprinus carpio)
transcriptome during environmental stress
Unpublished (2003)
Contact: Andrew R. Cossins
Laboratory for Environmental Gene Regulation
University of Liverpool
School of Biological Sciences, The Biosciences Building, Crown
Street, Liverpool, United Kingdom, L69 7ZE
Tel: +44(0)151-795-4510
Fax: +44(0)151-795-4431
Email: cossins@liv.ac.uk
Vector has been trimmed from this EST.
Plate: 10 row: d column: 20
Seq primer: Triplex 5' LD (5'-CTCGGAAGCGCGCATTTGTGTGT-3')
High quality sequence start: 30
High quality sequence stop: 568.
Location/Qualifiers
1..947
/organism="Cyprinus carpio"
/mol_type="mRNA"
/db_xref="taxon:7962"
/clone="10d20"
/sex="Male & female"
/tissue_type="Liver"
/dev_stage="Adult"
/lab_host="E.coli Electromax DH10B"
/clone_lib="Carp liver library 3"
/note="Vector: pTriplex2; Site 1: Sfil GGCCATTACGGCC;
Site 2: Sfil GGCGCTCGGCC; Normalized and serially
subtracted cDNA library prepared from liver of warm, cold
and hypoxia challenged animals"

ORIGIN
Alignment Scores:
Pred. No.: 1.31e-97 Length: 947
Score: 939.00 Matches: 173
Percent Similarity: 84.19% Conservative: 40
Best Local Similarity: 68.38% Mismatches: 40
Query Match: 70.23% Indels: 0
DB: 14 Gaps: 0

US-10-063-735-128 (1-260) x CA964199 (1-947)

QY 6 MetGluArgTrrArgAspArgLeuAlaLeuValThrGlyAlaSerGlyGlyIleGlyAla 25
Db 58 ATGGATCGCTGGAGAGTGGGTGCTCTGTCTACCTGGAGCTTCAGTGGGAATCGGAGCT 117
QY 26 AlaValAlaArgAlaLeuValGlnGlnGlyLeuLysValValGlyCysAlaArgThrVal 45
Db 118 GCAATCGCAAGTCTCTTGTCCAGCATGCGATGAGGTGATCGCTGCGCCAGAAATGTG 177
QY 46 GlyAsnIleGluGluLeuAlaGluCysLysSerAlaGlyTyrProGlyThrLeuIle 65
Db 178 GAGCAAAATTCAGAAATCGCAGCAGAAATGTGTGAGTGGTGGTATGAGCGCATCTGATC 237
QY 66 ProTyrArgCysAspLeuSerAsnGluGluAspIleLeuSerMetPheSerAlaIleArg 85
Db 238 CGGTACAAATGTGATCTGTCTGAGGAGCAGCATGATATCATGTTCTCTGGATCAAA 297
QY 86 SerGlnHisSerGlyValAspIleCysIleAsnAlaGlyLeuAlaArgProAspThr 105
Db 298 GTTCAACATCAGCGCATTCAGCTGTGCAATTAATGCTGGTTGGCTCTCCAGAGCCT 357
QY 106 LeuLeuSerGlySerThrSerGlyTrrLysAspMetPheAsnValAsnValLeuAlaLeu 125
Db 358 CTACTGAGTGGCAAACTAGTGGCTGGAGGACTATGATGATGTAATGTAATGGCCTG 417
QY 126 SerIleCysThrArgGluAlaTyrGlnSerMetLysGluArgAsnValAspAspGlyHis 145
Db 418 TCAGTGTGCACCGCTGAGGCTTACGCTCCATGAAAGAAAGAAATGATGATGTCAT 477
QY 146 IleIleAsnIleAsnSerMetSerGlyHisArgValLeuProLeuSerValThrHisPhe 165
Db 478 ATCAATTAATATCAACAGTATTTGTGACACCGGGTCTCAACCATGCTGATGCACACTTC 537
QY 166 TyrSerAlaThrLysTyrAlaValThrAlaLeuThrGluGlyLeuArgGlnGluLeuArg 185
Db 538 TACAGTGCAGCAAAATATGACGTGACGCTCTCACAGAAAGTTTGAGGCAAGATTACGA 597
QY 186 GluAlaGlnThrHisIleArgAlaThrCysIleSerProGlyValValGluThrGlnPhe 205
Db 598 GAGGCCAAACCCACATAGTCCACATATTTCTCTGGTTAGTGAGACAGCAATTT 657
QY 206 AlaPheLysLeuHisAspLysAspProGluLysAlaAlaAlaThrTyrGluGlnMetLys 225
Db 658 GCCTACCGACTCTTTAGCGANAACCCAGAAATAGCTGTGCTACCTACACNAGTGTAAAG 717
QY 226 CysLeuLysProGluAspValAlaGluAlaValIleTyrValLeuSerThrProAlaHis 245
Db 718 TGCCTACNAGCAGTTGACATCGCAAACTCAGTGGTGTATGCTCCTGAGTCTCTCTCTCAT 777
QY 246 IleGlnIleGlyAspIleGlnMetArgProThrGluGln 258
Db 778 GNTCAATTTGGGTGACATTGAGATCACCAGCTGGAGCAG 816

RESULT 19
LOCUS CA969358
DEFINITION Cctk06a22g15f1 Carp mixed tissue library 3 Cyprinus carpio cDNA
ACCESSION CA969358
VERSION 1
KEYWORDS EST.
SOURCE Cyprinus carpio (common carp)
ORGANISM Cyprinus carpio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Cyprinus.
1 (bases 1 to 819)
AUTHORS Gracey A.Y., Fraser E., Li W. and Cossins A.R.
TITLE Microarray and EST analysis of the carp (Cyprinus carpio)
JOURNAL transcriptome during environmental stress
COMMENT Unpublished (2003)
Contact: Andrew R. Cossins
Laboratory for Environmental Gene Regulation

University of Liverpool
School of Biological Sciences, The Biosciences Building, Crown
Street, Liverpool, United Kingdom, L69 7ZB
Tel: +44 (0)151-795-4510
Fax: +44 (0)151-795-4431
Email: cossins@liv.ac.uk
Vector has been trimmed from this EST.
Plate: 22 row: 9 column: 15
Seq primer: TripLex 5' LD (5'-CTCGGGAAGCGCCATTTGTTGGT-3')
High quality sequence start: 65
High quality sequence stop: 569.
Location/Qualifiers
1..819
/organism="Cyprinus carpio"
/mol_type="mRNA"
/db_xref="taxon:7962"
/clone="22g15"
/sex="Male & female"
/tissue_type="Skeletal white muscle, cardiac muscle,
kidney, brain, gill, intestinal mucosa"
/dev_stage="Adult"
/lab_host="E.coli Electromax DH10B"
/clone_lib="Carp mixed tissue library 3"
/note="Vector: pTriplex2; Site_1: Sfil GGCCATTACGGCC;
Site_2: Sfil GGCCGCTCGGCC; Normalized and serially
subtracted cDNA library prepared from mixed tissues of
warm, cold and hypoxia challenged animals"

ORIGIN

Alignment Scores:
Pred. No.: 2,96e-97 Length: 819
Score: 935.00 Matches: 173
Percent Similarity: 84.92% Conservative: 41
Best Local Similarity: 68.65% Mismatches: 38
Query Match: 69.93% Indels: 0
DB: 14 Gaps: 0
US-10-063-735-128 (1-260) x CA969358 (1-819)
QY 6 MetGluArgTrrArgAspArgLeuAlaLeuValThrGlyAlaSerGlyGlyIleGlyAla 25
Db 59 ATGGTCCGCTGAAAGGCGAGGCTTGTCTTGTCACTGGAGCTTCAGTGGGAATCGGAGCT 118
QY 26 AlaValAlaArgAlaLeuValGlnGlnGlyLeuLysValValGlyCysAlaArgThrVal 45
Db 119 GCAATCCCAAGTCTCTTGTCCAGCATGCGATGAAAGTGGGGGCTGTGCCAAATGTG 178
QY 46 GlyAsnIleGluGluLeuAlaGluCysLysSerAlaGlyTyrProGlyThrLeuIle 65
Db 179 GGGCAATTAAGAAATTCGCTGCTGAATGTGTAGTAGTGAATCAGCGGCATCTGATA 238
QY 66 ProTyrArgCysAspLeuSerAsnGluGluAspIleLeuSerMetPheSerAlaIleArg 85
Db 239 CCATATAAATGTGATCTGTCTGTAAGGACGACATGTTATCCATGTTCTCTCGGATCAAA 298
QY 86 SerGlnHisSerGlyValAspIleCysIleAsnAlaGlyLeuAlaArgProAspThr 105
Db 299 GTTCAACATCAGGGAGTTGAGTGTGCAATTAATATGCTGGTGTGGCTCTCCCTCGCCT 358
QY 106 LeuLeuSerGlySerThrSerGlyTrrLysAspMetPheAsnValAsnValLeuAlaLeu 125
Db 359 CTATTGAGCGGGACCCACCTGGCTGGAGGACTATGATGAGTGAATGTATTGGGCTG 418
QY 126 SerIleCysThrArgGluAlaTyrGlnSerMetLysGluArgAsnValAspAspGlyHis 145
Db 419 TCAGTGTGACCCCGTGAGGCATACCCTCCATGAAGAACGAAATATTGATGATGGTCTAT 478
QY 146 IleIleAsnIleAsnSerMetSerGlyHisArgValLeuProLeuSerValThrHisPhe 165
Db 479 ATCATTAAATTAACAGCATGAGTGGACACCGGGTCTGCAACAAATGCCGATGACACTTC 538
QY 166 TyrSerAlaThrLysTyrAlaValThrAlaLeuThrGluGlyLeuArgGlnGluLeuArg 185


```

source
1. .566
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S1SNU5s1-4-C01"
/sex="F"
/tissue_type="Ascites"
/cell_type="Lymphoblast-like"
/lab_host="T0p10F"
/clone_lib="S1SNU5s1"
/note="Organ: Stomach; Vector: pCNS; Site 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tabacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dt-selected mRNA by
priming with dr-tailed vector. The dr-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library. After analyzing and
sequencing about 2,000 ~ 3,000 colonies in original cDNA
library, the abundant cDNAs were selected and amplified by
PCR reaction using vector region primer including T7
promotor as 5' primer and N(dT)14 as 3' primer. The PCR
products were used as template for synthesis of
biotinylated single stranded RNA by in vitro transcription
reaction. The synthesized RNA probes were hybridized with
antisense single stranded cDNAs prepared from original
library and incubated with avidin-gel. After removing
DNA-RNA hybrids by centrifuge, the substracted cDNA
libraries were constructed by transformatioin of the
remaining DNA into competent cells E. coli Top10F' with
electroporation method."

ORIGIN
Alignment Scores:
Pred. No.: 1,31e-96 Length: 566
Score: 927.00 Matches: 180
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 69.33% Indels: 0
DB: 12 Gaps: 0

US-10-063-735-128 (1-260) x BM744024 (1-566)

QY 81 PheSerAlaIleArgSerGlnHisSerGlyValAspIleCysIleAsnAlaGlyLeu 100
Db 3 TCTCAGGTATCCGTTCTCAGCACACGCGGTAGACATCTGCATCAACAATCGTGGCTTG 62
QY 101 AlaArgProAspThrLeuLeuSerGlySerThrSerGlyTrpLysAspMetPheAsnVal 120
Db 63 GCCCGCGCTGACACCTGCTCTCAGCAGCACCGAGTGGTTGGAGGACATGTCATGTG 122
QY 121 AsnValLeuAlaLeuSerIleCysThrArgGluAlaTyrGlnSerMetIysGluArgAsn 140
Db 123 AACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCTACCATCTCCATGAGAGCGGAAT 182
QY 141 ValAspAspGlyHisIleIleAsnIleAsnSerMetSerGlyHisArgValLeuProLeu 160
Db 183 GTGGACGATGGGCACATCATTAACATCAATGCAATGCTCTGGCCACCGAGTGTACCCCTG 242
QY 161 SerValThrHisPheThrSerAlaThrIysTyrAlaValThrAlaIleuThrGluGlyLeu 180
Db 243 TCTGTGACCCACTTCTATAGTCCCAACCAAGTAGTCCGTCAGTCGCTGACAGGGAATG 302
QY 181 ArgGlnGluLeuArgGluAlaGlnThrHisIleArgAlaThrCysIleSerProGlyVal 200

```

303 AGGCAGAGCTTCGGAGGCCCGACCCACATCCGAGCCACGTCATCTCTCCAGGTGTG 362
201 ValGluThrGlnPheAlaPheIysLeuHisAspLysAspProGluLysAlaAlaIaThr 220
363 GTGGAGACACAATTCGCTTCAAACTCCACGACACAGACCCCTGAGAAGCGAGTGCACC 422
221 TyrGluGlnMetLysCysLeuLysProGluAspValAlaGluAlaValIleTyrValLeu 240
423 TATGACCAATGAAGTGTCTCAAAACCCGAGGATGTGCGCCAGGCTGTATCTACGTCCTC 482
241 SerThrProAlaHisIleGlnIleGlyAspIleGlnMetArgProThrGluGlnValThr 260
483 AGCACCCCGCACACATCCAGATTGGAGATCCAGATGAGGCCACCGAGAGGTGACC 542

BM742038 573 bp mRNA linear EST 01-MAR-2002
K-EST0014773 S1SNU5s1 Homo sapiens cDNA clone S1SNU5s1-1-B09 5',
mRNA sequence.
BM742038
BM742038.1 GI:19063367
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 573)
Kim.N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
Unpublished (2002)
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-339, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 1 row: B column: 09
High quality sequence stop: 573.
Location/Qualifiers
1. .573
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S1SNU5s1-1-B09"
/sex="F"
/tissue_type="Ascites"
/cell_type="Lymphoblast-like"
/lab_host="T0p10F"
/clone_lib="S1SNU5s1"
/note="Organ: Stomach; Vector: pCNS; Site 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tabacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dt-selected mRNA by
priming with dr-tailed vector. The dr-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library. After analyzing and
sequencing about 2,000 ~ 3,000 colonies in original cDNA
library, the abundant cDNAs were selected and amplified by
PCR reaction using vector region primer including T7
promotor as 5' primer and N(dT)14 as 3' primer. The PCR
products were used as template for synthesis of
biotinylated single stranded RNA by in vitro transcription
reaction. The synthesized RNA probes were hybridized with
antisense single stranded cDNAs prepared from original
library and incubated with avidin-gel. After removing
DNA-RNA hybrids by centrifuge, the substracted cDNA
libraries were constructed by transformatioin of the
remaining DNA into competent cells E. coli Top10F' with
electroporation method."

promotor as 5' primer and N(dT)14 as 3' primer. The PCR products were used as template for synthesis of biotinylated single stranded RNA by in vitro transcription reaction. The synthesized RNA probes were hybridized with antisense single stranded cDNAs prepared from original library and incubated with avidin-gel. After removing DNA-RNA hybrids by centrifuge, the subtraced cDNA libraries were constructed by transformation of the remaining DNA into competent cells *E. coli* Top10f' with electroporation method."

ORIGIN

Alignment Scores:
Pred. No.: 1,34e-96 Length: 573
Score: 927.00 Matches: 180
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 69.33% Indels: 0
DB: 12 Gaps: 0

US-10-063-735-128 (1-260) x BM742038 (1-573)

Qy 81 PheSerAlaIleArgSerGlnHisSerGlyValAspIleCysIleAsnAsnAlaGlyLeu 100
Db 3 TTCTCAGCTATCGTCTCTCAGCACAGCGGTGTAGACATCTGCATCAACATGCTGGCTTG 62
Qy 101 AlaArgProAspThrLeuLeuSerGlySerThrSerGlyTrpIysAspMetPheAsnVal 120
Db 63 GCCCGCCTTGACACCTGCTCTCAGGACAGCAGTGGTGGGAAGACATGTTCAATGTG 122
Qy 121 AsnValLeuAlaLeuSerIleCysThrArgGluAlaTyrGlnSerMetIysGluArgAsn 140
Db 123 AAGTCTGGCCCTCAGCATCTGCACACGGGAAGCCTACCATGCCATGAAGAGCGGAAT 182
Qy 141 ValAspAspGlyHisIleIleAsnIleAsnSerMetSerGlyHisArgValLeuProLeu 160
Db 183 GTGGACGATGGGCACATCATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTG 242
Qy 161 SerValThrHisPheTyrSerAlaThrLysTyrAlaValThrAlaLeuThrGluGlyLeu 180
Db 243 TCTGTGACCCACTTCTATAGTGCACCAAGTATGCCGTCTCACTCGCTGCAGAGGACTG 302
Qy 181 ArgGlnGluLeuArgGluAlaGlnThrHisIleArgAlaThrCysIleSerProGlyVal 200
Db 303 AGCGAAGACTTCGGGAGGCCACAGACCAATCCGAGCCACGTGCATCTCTCAGGTGTG 362
Qy 201 ValGluThrGlnPheAlaPheLysLeuHisAspLysAspProGluLysAlaAlaIleThr 220
Db 363 GTGGAGACACAAATTCGCTTCAAACTCCAGCACAGGACCCCTGAGAAGGCGAGTGCACC 422
Qy 221 TyrGluGlnMetLysCysLeuLysProGluAspValAlaGluAlaValIleTyrValLeu 240
Db 423 TATGACAAATGAAGTGTCTCAAAACCGAGGATGTGGCCGAGGCTGTATCTACGTCTC 482
Qy 241 SerThrProAlaHisIleGlnIleGlyAspIleGlnMetArgProThrGluGlnValThr 260
Db 483 AGCACCCCGCACACATCCAGATGGAGATCCAGATGAGGCCCCACGAGGAGGTGACC 542

RESULT 23

BM764342
LOCUS K-EST0045842 S1SNU5 Homo sapiens cDNA clone S1SNU5-39-A06 5', mRNA
DEFINITION sequence.
EST.

BM764342 GI:119093957

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 602)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and

TITLE
JOURNAL
COMMENT

Kim,Y.S.
21C Frontier Korean EST Project 2001
Unpublished (2002)
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 39 row: A column: 06
High quality sequence stop: 602.

FEATURES

source

1..602
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S1SNU5-39-A06"
/sex="F"
/tissue_type="Ascites"
/cell_type="Lymphoblast-like"
/cell_line="SNU-5"
/lab_host="Top10f"
/clone_lib="S1SNU5"
/note="Organ: Stomach; Vector: pCNS; Site 1: EcoRI;
Site 2: NotI; The poly (A) + RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dT-selected mRNA by
priming with dT-tailed vector. The dT-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with *E. coli* DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells *E. coli* Top10f' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

ORIGIN

Alignment Scores:
Pred. No.: 1,46e-96 Length: 602
Score: 927.00 Matches: 180
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 69.33% Indels: 0
DB: 12 Gaps: 0

US-10-063-735-128 (1-260) x BM764342 (1-602)

Qy 81 PheSerAlaIleArgSerGlnHisSerGlyValAspIleCysIleAsnAsnAlaGlyLeu 100
Db 3 TTCTCAGCTATCGTCTCTCAGCACAGCGGTGTAGACATCTGCATCAACATGCTGGCTTG 62
Qy 101 AlaArgProAspThrLeuLeuSerGlySerThrSerGlyTrpIysAspMetPheAsnVal 120
Db 63 GCCCGCCTTGACACCTGCTCTCAGGACAGCAGTGGTGGGAAGACATGTTCAATGTG 122
Qy 121 AsnValLeuAlaLeuSerIleCysThrArgGluAlaTyrGlnSerMetIysGluArgAsn 140
Db 123 AAGTCTGGCCCTCAGCATCTGCACACGGGAAGCCTACCATGCCATGAAGAGCGGAAT 182
Qy 141 ValAspAspGlyHisIleIleAsnIleAsnSerMetSerGlyHisArgValLeuProLeu 160
Db 183 GTGGACGATGGGCACATCATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTG 242
Qy 161 SerValThrHisPheTyrSerAlaThrLysTyrAlaValThrAlaLeuThrGluGlyLeu 180
Db 243 TCTGTGACCCACTTCTATAGTGCACCAAGTATGCCGTCTCACTCGCTGCAGAGGACTG 302
Qy 181 ArgGlnGluLeuArgGluAlaGlnThrHisIleArgAlaThrCysIleSerProGlyVal 200

```

303 AGCAAGAGCTTCGGAGGCCAGACCCACATCCAGCCAGCTGTCATCTCTCCAGGTGTG 362
201 ValGluThrGlnPheAlaPheLysLeuHisAspLysProGluLysAlaAlaThr 220
363 GTGGAGACAAATTCGCTTCAAACTCCAGCAAGACCTTGAGAGGAGCGAGTGCACC 422
221 TyrGluGlnMetLysCysLeuLysProGluAspValAlaGluAlaValIleTyrValLeu 240
423 TATGAGCAATGAAGTGTCTCAACCGAGAGTGGCCGAGGCTGTATCTACGTCCTC 482
241 SerThrProAlaHisIleGlnIleGlyAspIleGlnMetArgProThrGluGlnValThr 260
483 AGACCCCGCCGACATCCAGATTGGAGACATCCAGATGAGGCCACGAGGAGGTGACC 542

RESULT 24
BM755987
LOCUS
DEFINITION
K-EST0034121 S1SNUS Homo sapiens cDNA clone S1SNUS-38-A10 5', mRNA
sequence.
BM755987
VERSION
BM755987.1 GI:19085602
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 640)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
Unpublished (2002)
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 38 row: A column: 10
High quality sequence stop: 640.
Location/Qualifiers
1..640
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S1SNUS-38-A10"
/sex="F"
/tissue_type="Ascites"
/cell_type="Lymphoblast-like"
/lab_host="T0p10F"
/clone_lib="S1SNUS"
/note="Organ: Stomach; Vector: pcNS; Site 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tabacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transfection of
competent cells E. coli Top10F, by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

FEATURES
source
1..640
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S1SNUS-38-A10"
/sex="F"
/tissue_type="Ascites"
/cell_type="Lymphoblast-like"
/lab_host="T0p10F"
/clone_lib="S1SNUS"
/note="Organ: Stomach; Vector: pcNS; Site 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tabacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transfection of
competent cells E. coli Top10F, by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

ORIGIN
Alignment Scores:

```

```

Pred. No.: 1.62e-96 Length: 640
Score: 927.00 Matches: 180
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 69.33% Indels: 0
DB: 12 Gaps: 0

US-10-063-735-128 (1-260) x BM755987 (1-640)

QY 81 PheSerAlaIleArgSerGlnHisSerGlyValAspIleCysIleAsnAlaGlyLeu 100
Db 3 TTCTGAGTATCGTCTTCAGCACAGCGGTGTAGACATCTGCATCAACAATGCTGGCTTG 62
QY 101 AlaArgProAspThrLeuLeuSerGlySerThrSerGlyTrpLysAspMetPheAsnVal 120
Db 63 GCCCGCCTGCACACCCCTGCTCTCAGGCAGCACCCAGTGGTGGAGGACATGTTCAATGTG 122
QY 121 AsnValLeuAlaLeuSerIleCysThrArgGluAlaTyrGlnSerMetLysGluArgAsn 140
Db 123 AAGTCTGGCCCTCAGCATCTGCACACGGGAAGCCTACCACTCCATGAAGGAGCGCAAT 182
QY 141 ValAspAspGlyHisIleIleAsnIleAsnSerMetSerGlyHisArgValLeuProLeu 160
Db 183 GTGGAGCATGGGCACATCATTAACATCAATGATGCTGGCCACCGAGTGTACCCCTG 242
QY 161 SerValThrHisPheTyrSerAlaThrLysTyrAlaValThrAlaLeuThrGluGlyLeu 180
Db 243 TCTGTGACCACTTCTATAGTCCCAAGTATGCCGTCACTCGCTGACAGAGGAGCTG 302
QY 181 ArgGlnGluLeuArgGluAlaGlnThrHisIleArgAlaThrCysIleSerProGlyVal 200
Db 303 AGCGAAGAGCTTCGGAGGCCAGACCCACATCCGAGCCACGTCGATCTCTCCAGGTGTG 362
QY 201 ValGluThrGlnPheAlaPheLysLeuHisAspLysAspProGluLysAlaAlaThr 220
Db 363 GTGGAGACAAATTCGCTTCAAACTCCAGCAAGACCCCTGAGAGGAGCTGCCACC 422
QY 221 TyrGluGlnMetLysCysLeuLysProGluAspValAlaGluAlaValIleTyrValLeu 240
Db 423 TATGAGCAATGAAGTGTCTCAACCGAGGATGTGGCCGAGGCTGTATCTACGTCCTC 482
QY 241 SerThrProAlaHisIleGlnIleGlyAspIleGlnMetArgProThrGluGlnValThr 260
Db 483 AGACCCCGCCGACATCCAGATTGGAGACATCCAGATGAGGCCACGAGGAGGTGACC 542

RESULT 25
BI391549
LOCUS
DEFINITION
pgpln.pk004.m18 Normalized Chicken Pituitary/Hypothalamus/Pineal
Library Gallus gallus cDNA clone pgpln.pk004.m18 5', similar to
gi|13652714 ref|XP_018530.1| hypothetical protein MGC4172 [Homo
sapiens], mRNA sequence.
BI391549
ACCESSION
BI391549.1 GI:15084831
VERSION
BI391549.1
KEYWORDS
SOURCE
ORGANISM
Gallus gallus (chicken)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 678)
Porter,T.E. and Cogburn,L.A.
ESTs from Normalized Chicken Pituitary/Hypothalamus/Pineal cDNA
Library USDA/IFAFS Animal Genome Project
Unpublished (2001)
Contact: Larry A. Cogburn
University of Delaware
Townsend Hall, Newark, DE 19717, USA
Tel: 302-831-1335
Fax: 302-831-2822
Email: cogburn@udel.edu, www.chickest.udel.edu.
Location/Qualifiers
1..678

```

```

/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Commercial broiler chicken"
/db_xref="taxon:9031"
/clone="pgpin.pk004.ml8"
/sex="Male and Female"
/tissue_type="Pituitary Gland/Hypothalamus/Pineal Gland"
/dev_stage="Embryonic (dl2,dl4,dl9); post-hatch
(w1,w3,w5,w7,w9)"
/lab_host="E. Coli EMDH10B"
/clone_lib="Normalized Chicken
Pituitary/Hypothalamus/Pineal Library"
/notes="Vector: pCMVSPORT6; Library made from equivalent
pools of total RNA isolated from each tissue at different
ages. Single pass sequencing from 5'-end"

```

ORIGIN

Alignment Scores:
 Pred. No.: 1.8e-96 Length: 678
 Score: 927.00 Matches: 175
 Percent Similarity: 88.74% Conservative: 22
 Best Local Similarity: 78.83% Mismatches: 25
 Query Match: 69.33% Indels: 0
 DB: 12 Gaps: 0

US-10-063-735-128 (1-260) x BI391549 (1-678)

```

QY 6 MetGluArgTTPArgAspArgLeuAlaLeuValThrGlyAlaSerGlyGlyIleGlyAla 25
DB 13 ATGGACGGTGGACCGGGCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 72
QY 26 AlaValAlaArgAlaLeuValGlnGlnGlyLeuLeuValGlyCysAlaArgThrVal 45
DB 73 GCGTGGCGCGGGCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 132
QY 46 GlyAsnIleGluLeuAlaGluCysLeuSerAlaGlyTyrProGlyThrLeuIle 65
DB 133 GACAAGATCGAAGATTTGGGAGCTGATGCCAGTGGCGGCTGACCTGGAGCCCTCAT 192
QY 66 ProTyrArgCysAspLeuSerAsnGluGluAspIleLeuSerMetPheSerAlaIleArg 85
DB 193 CCGTACAGTGTGATCTCTCCATGAAGAGAGATCTGTCCATGTCTCTGCGATCAAG 252
QY 86 SerGlnHisSerGlyValAspIleCysIleAsnAlaGlyLeuAlaArgProAspThr 105
DB 253 ACCCTTCACCGAGGAGTGTATGTGTGCATCAACACGAGGGCTGGCTGCCCGGAGCCT 312
QY 106 LeuLeuSerGlySerThrSerGlyTyrLysAspMetPheAsnValAsnValLeuAlaLeu 125
DB 313 CTGCTCTCAGGGAAGACAGAGGCTGGCGACATGTAGACGTCAAGTGTGGCTGTG 372
QY 126 SerIleCysThrArgGluAlaTyrGlnSerMetLysGluArgAsnValAspGlyHis 145
DB 373 AGCATCTGCCACCGAGAGGCTATCATCTCATGAAGGAGAGAAACATCGATGACGGTCA 432
QY 146 IleIleAsnIleAsnSerMetSerGlyHisArgValLeuProLeuSerValThrHisPhe 165
DB 433 ATTATTAAACATTAAACAGCATGAATGGTTCACAGCGTGTGGCCACAGTGGTGGTCAATTT 492
QY 166 TyrSerAlaThrLysTyrAlaValThrAlaLeuThrGluGlyLeuArgGlnGluLeuArg 185
DB 493 TACAGTCCCAACCAAGTATGAGTCCAGGCGCTTCACGANNNGGTGAGCAGAACTCAGA 552
QY 186 GluAlaGlnThrHisIleArgAlaThrCysIleSerProGlyValValGluThrGlnPhe 205
DB 553 GAACAAAGACTCATATACGAGCTACATGTATATCTCCAGACTGTGTAGAAACGGGATTT 612
QY 206 AlaPheLysLeuHisAspLysAspProGluLysAlaAlaThrTyrGluGlnMetLys 225
DB 613 GCTTTTAAACTTCATGATTAATGATGACCTCTGCTGCAACCTATGNNNNCATTCGG 672
QY 226 CysLeu 227

```

```

Db 673 TGTCTC 678
RESULT 26
BU212233
LOCUS
DEFINITION
BU212233 849 bp mRNA linear EST 25-NOV-2002
603950353F1 CSEQCHN03 Gallus gallus cDNA clone CHEST909j14 5', mRNA
sequence.
BU212233
EST.
BU212233.1 GI:25386475
SOURCE
Gallus gallus (chicken)
ORGANISM
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 849)
Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
22335534
12445352
Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
1..849
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="CHEST909j14"
/tissue_type="whole embryo"
/dev_stage="20-21"
/lab_host="DH10B"
/clone_lib="CSEQCHN03"
/notes="Organ: whole embryo; Vector: pBluescript II KS(+);
Site 1: EcoRI; Site 2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following C in the first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

```

ORIGIN

Alignment Scores:
 Pred. No.: 2.25e-95 Length: 849
 Score: 919.00 Matches: 174
 Percent Similarity: 91.28% Conservative: 25
 Best Local Similarity: 79.82% Mismatches: 19
 Query Match: 68.74% Indels: 1
 DB: 13 Gaps: 0

US-10-063-735-128 (1-260) x BU212233 (1-849)

```

QY 43 ArgThrValGlyAsnIleGluGluLeuAlaGluCysLysSerAlaGlyTyrProGly 62
DB 6 CGCAGCGTGACCAAGATCGAGAATTCGACGATGATGCCAGATGCCGCTACCTGGG 65
QY 63 ThrieulleProTyrArgCysAspLeuSerAsnGluAlaSplLeuSerMetPheSer 82

```

Db 66 ACCTCATTCCTGACAGTGTGATCTCTCCAAATGAAGAGAGATCCTGCTCCATGTTCTCT 125

QY 83 AlAlaArgSerGlnHisSerGlyValAspIleCysIleAsnAlaGlyLeuAlaArg 102

Db 126 GCCATCAAGACCTTCCACA-GAGGTGATGTGTGTCATCAACACGCGGGTGGCTCGC 184

QY 103 ProAspThrLeuLeuSerGlySerThrSerGlyTyrLysAspMetPheAsnValAsnVal 122

Db 185 CCGGAGCCTCTGCTCTCAGGGAAGACAGAAAGCGTGGCGGACATGATGACGTCAACGTG 244

QY 123 LeuAlaLeuSerIleCysThrArgGluAlaTyrGlnSerMetLysGluArgAsnValAsp 142

Db 245 ATGGCTGTGAGCATCTGCCCGGAGAGCGCTATCAGTCCATGAAGAGAGAAACATCGAT 304

QY 143 AspGlyHisIleIleAsnIleAsnSerMetSerGlyHisArgValLeuProLeuSerVal 162

Db 305 GACGTCATATTATTAAACATTACACATGAATGGTCACAGCGTGTGCCAGTCGGTG 364

QY 163 ThrHisPheTyrSerAlaThrLysTyrAlaValThrAlaLeuThrGluGlyLeuArgGln 182

Db 365 GTGCATTTTACAGTGCACCAAGTATGCAGTCACGCGCTCACAGAGGGGCTGAGGCA 424

QY 183 GluLeuArgGluAlaGlnThrHisIleArgAlaThrCysIleSerProGlyValValGlu 202

Db 425 GAACTCAGAGAAGCAAGACTCATATACAGGTACATGTATATCTCCAGGACTGGTAGAA 484

QY 203 ThrGlnPheAlaPheLysLeuHisAspLysAspProGluLysAlaAlaAlaThrTyrGlu 222

Db 485 ACGGATTTGCTTTTAACTTCATGATATGACCTGAGAGAGCTGCTCAACACCTATGAG 544

QY 223 GlnMetLysCysLeuLysProGluAspValAlaGluAlaValIleTyrValLeuSerThr 242

Db 545 AGCATTCGGTGTCTCAAGCTCAAGATATGGCTAATGCTGTATATATGCTCTTAGTGCC 504

QY 243 ProAlaHisIleGlnIleGlyAspIleGlnMetArgProThrGluGlnValThr 260

Db 605 CCACCTCATGTACAGATTGGAGATATACAGATGAGGCGCCACGGAACAATATCA 658

RESULT 27
BX864140
LOCUS
DEFINITION BX864140 Rainbow trout multi-tissues (tcba) linear EST 16-DEC-2003
ACCESSION BX864140
VERSION BX864140
KEYWORDS BX864140.1 GI:39962970
SOURCE
ORGANISM Oncorhynchus mykiss (rainbow trout)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
1 (bases 1 to 743)
Govoroun,M., Guiguen,Y. and Le Gac,F.
Construction and primary characterization of normalized cDNA
libraries in rainbow trout, Oncorhynchus mykiss
Unpublished (2003)
Contact: Guiguen Y
INRA - SCRIBE
Campus de beaulieu, RENNES cedex, 35042, France
Tel: 02.23.48.50.09
Fax: 02.23.48.50.20
Email: Yann.Guiguen@beaulieu.rennes.inra.fr
Sequence cleaned of vector, adaptor and repetitions. Contact us
at sigenassupport@jouy.inra.fr to obtain the chromatogram of this
sequence.
Plate: 0017 row: e column: 9
Seq primer: M13R.
Location/Qualifiers
1. 743
/organism="Oncorhynchus mykiss"
/mol_type="mRNA"
/db_xref="taxon:8022"
/clones="tcba0017c.e.09"

FEATURES
source

CF595166
LOCUS
RESULT 28
CF595166
912 bp mRNA linear EST 08-OCT-2003

/tissue_type="multi-tissues"
/dev_stage="from embryos to adults"
/lab_host="DH10B"
/clone_lib="Rainbow trout multi-tissues (tcba)"
/note="Vector: pT73D-pac; Clone distribution : AGENAE
Resource centre: Francois PIUMI
Francois.Piumi@jouy.inra.fr, INRA, CEA Radiobiologie et
Etude du genome (JREG), Domaine de Vilvert, 78352,
Jouy-en-Josas cedex, FRANCE, +33 (0) 1.34.65.28.02, +33
(0) 1.34.65.22.73"

ORIGIN

Alignment Scores:
Pred. No.: 115e-94 Length: 743
Score: 912.00 Matches: 174
Percent Similarity: 86.81% Conservative: 30
Best Local Similarity: 74.04% Mismatches: 30
Query Match: 68.21% Indels: 1
DB: 13 Gaps: 0
US-10-063-735-128 (1-260) x BX864140 (1-743)
QY 6 MetGluArgTyrArgAspArgLeuAlaLeuValThrGlyAlaSerGlyGlyIleGlyAla 25
Db 41 ATGGATCGCTGGAAGGTAGAGTGGCGCTTGTACCGGGGCTTCTGTGGAAATCGAGCG 100
QY 26 AlaValAlaArgAlaLeuValGlnGlnGlyLeuLysValValGlyCysAlaArgThrVal 45
Db 101 GCGATTTCAGAGCTCTTGTCTCAGCAGCGCATGAAGTTGTCCGCTGTGCCAGGAACGTC 160
QY 46 GlyAsnIleGluGluLeuAlaAlaGluCysLysSerAlaGlyTyrProGlyThrLeuIle 65
Db 161 GAGAAATAGAGAAATCGCAGCAGAGTGTGAGCGCTGGGACAGTCGACCCCTTGT 220
QY 66 ProTyrArgCysAspLeuSerAsnGluGluAspIleLeuSerMetPheSerAlaIleArg 85
Db 221 CCTTACAATGTGACCTTTCAAATGAAGAGGAGATCCTTTCCATGTTCTCTGCCATCAAG 280
QY 86 SerGlnHisSerGlyValAspIleCysIleAsnAsnAlaGlyLeuAlaArgProAspThr 105
Db 281 ACTCTCCATCAGGAGTGGATGTGTGCATCAACAATGTGGCTGGCCCATTCAGAGTCA 340
QY 106 LeuLeuSerGlySerThrSerGlyTyrLysAspMetPheAsnValAsnValLeuAlaLeu 125
Db 341 CTGTGAATGGCAAAACCGATGGCTGGGACAAATGATTGATCTGAACGTCATTGCATTG 400
QY 126 SerIleCysThrArgGluAlaTyrGlnSerMetLysGluArgAsnValAspAspGlyHis 145
Db 401 TCTATCTGCACAGAGAGGCATACCGCTCATGAGGAGAGAGAGGTGGATGGACAT 460
QY 146 IleIleAsnIleAsnSerMetSerGlyHisArgValLeuProLeuSerValThrHisPhe 165
Db 461 ATCATCAACATAAACAGTATGAGTGGACATCGTGTAGTCTTTCAGTGCTGATACACATTC 520
QY 166 TyrSerAlaThrLysTyrAlaValThrAlaLeuThrGluGlyLeuArgGlnGluLeuArg 185
Db 521 TATAGTGTCTACCAAGTATGCTGTGACTGCCCTGACCGCAANGGTTGCCGTCAAGAGTCGA 580
QY 186 GluAlaGlnThrHisIleArgAlaThrCysIleSer-ProGlyValValGluThrGlnPh 205
Db 581 GAGGCTAAACCCACATTCGAGCCACCGTGTATATCTCCCTGGTATAGTGGAAACGGAAAT 640
QY 205 eAlaPheLysLeuHisAspLysAspProGluLysAlaAlaAlaThrTyrGluGlnMetLys 225
Db 641 TGCTTTCCGACTTCACAGCCTCCATCCAGAGAAGGCTGTGCTACCTACACAGATGAA 700
QY 225 sCysLeuLysProGluAspValAlaGluAlaValIleTyrVal 239
Db 701 GTGTTTGGAAAGCAGTCGACATTGCCAGTGTGTAACGATATGTC 743

```

DEFINITION  AGENCOURT_15714115 NIH_ZGC_8 Danio rerio cDNA clone IMAGE:7015515
5', mRNA sequence.
ACCESSION   CF595166
VERSION     CF595166.1  GI:36350329
KEYWORDS   EST.
SOURCE      Danio rerio (zebrafish)
ORGANISM    Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1 (bases 1 to 912)
            NIH-MGC http://mgc.nci.nih.gov/
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
AUTHORS     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Takuya Sakaguchi and Didier Stainier
cDNA Library Preparation: Dr. Sumio Sugano
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLMW14730 row: k column: 01
High quality sequence start: 25
High quality sequence stop: 691.
FEATURES             source
    source
        1..912
            /organism="Danio rerio"
            /mol_type="mRNA"
            /db_xref="taxon:7955"
            /clone="IMAGE:7015515"
            /lab_host="DH10B TonA"
            /clone_lib="NIH ZGC 8"
            /note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII,
            Site_2: DraIII"
ORIGIN
Alignment Scores:
Pred. No.:      4,78e-94      Length:      912
Score:          908.00      Matches:    169
Percent Similarity: 82.68%      Conservative: 41
Best Local Similarity: 66.54%      Mismatches: 44
Query Match:     67.91%      Indels:     0
DB:              14          Gaps:         0

US-10-063-735-128 (1-260) x CF595166 (1-912)
QY      6 MetGluArgTrpArgAspArgLeuAlaLeuValThrGlyAlaSerGlyGlyLeuGlyAla 25
Db      59 ATGGATCGCTGGAGAGCGCAGAGTGTCTTGTCTCAGTGGAGCTTCAGTGGGAATAGGAGCT 118
QY      26 AlaValAlaArgAlaLeuValGlnGlnGlyLeuLysValValGlyCysAlaArgThrVal 45
Db      119 GCAATCGCAAAATCTCTTGTCCAGCATGGATGAGTGGATGCGATGCCAGAAACGGT 178
QY      46 GlyAsnIleGluGluLeuAlaAlaGluCysLysSerAlaGlyTyrProGlyThrLeuIle 65
Db      179 GAACAAATTAAAGAACTGGCGCTGAATGTGTAGTGGATACAGCGGTCTGTGTTT 238
QY      66 ProTyrArgCysAspLeuSerAsnGluLysAlaPheLeuSerMetPheSerAlaIleArg 85
Db      239 CCATATAAATGTGATCTTTTCAGTAGAGGATGAGTGTGTCCATGTGTCTCTGGATCAAG 298
QY      86 SerGlnHisSerGlyValApeIleCysIleAsnAsnAlaGlyLeuAlaArgProAspThr 105
Db      299 GCTCAACATAAGGTTGTGATGTGTCATTAATTAATGCTGTGTTTAGCTCTGCCAGAGCCT 358
QY      106 LeuLeuSerGlySerThrSerGlyTrpLysAspMetPheAsnValLeuAlaLeu 125

```

```

Db      359 CTGTTAGTGGCAAAACCAGCAGCTGGAGGACTATGATGGAGCTCAATGTTCATGGCCCTG 418
QY      126 SerIleCysThrArgGluAlaTyrGlnSerMetLysGluArgAsnValAspAspGlyHis 145
Db      419 GCAGTGTGCACACGCTGAGGCTTACAGTCCATGAAGGAAAGAAAGTGAATGACGCCCAT 478
QY      146 IleIleAsnIleAsnSerMetSerGlyHisArgValLeuProLeuSerValThrHisPhe 165
Db      479 ATCAATATATATACAGTATTGTGGACACCGTGTGCTTAACCTATGCTGATGACACATTC 538
QY      166 TyrSerAlaThrLysTyrAlaValThrAlaLeuThrGluGlyLeuArgGlnGluLeuArg 185
Db      539 TACACCGCTACTAATACGCACTGACCGCCCTCACAGAAGGTTTCAGACACAGATTACGA 598
QY      186 GluAlaGlnThrHisIleArgAlaThrCysIleSerProGlyValValGlnThrGlnPhe 205
Db      599 GAGGCCAAGACACACATAGTCCACAGGAATATCTCTGTGTATAGTGAAGACAGAATTT 658
QY      206 AlaPheLysLeuHisAspLysAspProGluLysAlaAlaAlaThrTyrGluGlnMetLys 225
Db      659 GCCTACCGCACTTTTAGTGTATGACCAAGAAAGGCTGCTGTATGTAATAATAGTGTGAG 718
QY      226 CysLeuLysProGluAspValAlaGluAlaValIleTyrValLeuSerThrProAlaHis 245
Db      719 TGCCTGCAAGCTGATGACATAACCCCAAGCAGTGTGTATGTTCTTAAGTGTCTCTCTCAC 778
QY      246 IleGlnIleGlyAspIleGlnMetArgProThrGluGlnVal 259
Db      779 GTCCAAATCGGTGACCTCGAGTTAAACGCTGTGGGACCGATC 820

RESULT 29
BM848012      627 bp      mRNA      linear      EST 06-MAR-2002
LOCUS        K-EST0127637 S13KMS5 Homo sapiens cDNA clone S13KMS5-43-A10 5',
DEFINITION   mRNA sequence.
ACCESSION    BM848012
VERSION       BM848012.1  GI:19204411
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
REFERENCE     1 (bases 1 to 627)
AUTHORS       Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
              Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
              Kim,Y.S.
              21C Frontier Korean EST Project 2001
              Unpublished (2002)
              Contact: Kim YS
              Genome Research Center
              Korea Research Institute of Bioscience & Biotechnology
              52 Soeun-dong Yuseong-gu, Daejeon 305-333, South Korea
              Tel: +82-42-860-4470
              Fax: +82-42-860-4409
              Email: yongsung@mail.kribb.re.kr
              Plate: 43 row: A column: 10
              High quality sequence stop: 627.
FEATURES     Location/Qualifiers
              1..627
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="S13KMS5-43-A10"
              /tissue_type="myeloma"
              /cell_line="KMS-5"
              /lab_host="Top10F"
              /clone_lib="S13KMS5"
              /note="Vector: pCNS; Site 1: EcoRI; Site 2: NotI; The poly
              (A)+ RNA was dephosphorylated with bacterial alkaline
              phosphatase (BAP) and then decapped with tabacco acid
              pyrophosphatase (TAP). The decapped intact mRNA was
              ligated with DNA-RNA linker including EcoR I site by
              treatment of T4 RNA ligase and the first strand cDNA was

```

synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10⁺ by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORIGIN

Alignment Scores:
 Pred. No.: 5,52e-94 Length: 627
 Score: 905.00 Matches: 176
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 67.69% Indels: 0
 DB: 12 Gaps: 0

US-10-063-735-128 (1-260) x BM848012 (1-627)

QY 1 MetAlaArgProGlyMetClnuArgTrpArgAspArgLeuAlaLeuValThrGlyAlaSer 20
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 21 GlyGlyIleGlyAlaAlaValAlaArgAlaLeuValGlnGlnGlyLeuLysValValGly 40
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 158 GGGGGCATCG 217
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 41 CysAlaArgThrValGlyAsnIleGluGluLeuAlaAlaGluCysLysSerAlaGlyTyr 60
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 61 ProGlyThrLeuIleProTyrArgCysAspLeuSerAsnGluGluAspIleLeuSerMet 80
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 278 CCCGGGACTTTGATCCCTACAGATGTGACCTATCAAAATGAAGAGACATCCTCTCCATG 337
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 81 PheSerAlaIleArgSerGlnHisSerGlyValAspIleCysIleAsnAlaGlyLeu 100
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 338 TTCTCAGCTATCCGTTCTCAGCAGCGGTGAGACATCTGCATCAACATGCTGGCTG 397
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 101 AlaArgProAspThrLeuLeuSerGlySerThrSerGlyTrpLysAspMetPheAsnVal 120
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 398 GCCCGGCTGACACCTGCTCTCAGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 457
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 121 AsnValLeuAlaLeuSerIleCysThrArgGluAlaTyrGlnSerMetLysGluArgAsn 140
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 458 AACGTGCTGGGCCCTCAGCATCTGCACACGGGAAGCCTACCACTCCATGAAGGAGCGGAAT 517
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 141 ValAspAspGlyHisIleAsnIleAsnSerMetSerGlyHisArgValLeuProLeu 160
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 518 GTGGACGATGGGCACATCATTAACATCAATGATGCTGTGCCACCGAGTGTACCCCTG 577
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 161 SerValThrHisPheTyrSerAlaThrLysTyrAlaValThrAlaLeu 176
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 578 TCTGTGACCCACTTCTATAGTGCACCAAGTATGCCGTCACCTGCGCTG 625

RESULT 30

BM737591

LOCUS

DEFINITION K-EST0000094 S1SNU5 Homo sapiens cDNA clone S1SNU5-21-B02 5', mRNA

sequence.

VERSION BM737591

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS Kim,N.S., Hahn,X., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and

TITLE
JOURNAL
COMMENT

Kim,Y.S.
 21C Frontier Korean EST Project 2001
 Unpublished (2002)
 Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 21 row: B column: 02
 High quality sequence stop: 582.

FEATURES

source

1..582
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S1SNU5-21-B02"
 /sex="F"
 /tissue_type="Ascites"
 /cell_type="Lymphoblast-like"
 /lab_host="Top10F"
 /clone_lib="S1SNU5"
 /note="Organ: Stomach; Vector: pCNS; Site_1: EcoRI;
 Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
 bacterial alkaline phosphatase (BAP) and then decapped
 with tabacco acid pyrophosphatase (TAP). The decapped
 intact mRNA was ligated with DNA-RNA linker including EcoR
 I site by treatment of T4 RNA ligase and the first strand
 cDNA was synthesized from oligo dt-selected mRNA by
 priming with dt-tailed vector. The dt-tailed vector was
 adjusted to have about 60nt. The cDNA vector was
 circularized with E. coli DNA ligase after digestion of
 EcoRI which site is also included in vector. An RNA strand
 converted to a DNA strand by Okayama-Berg method. The
 obtained cDNA vectors were used for transformation of
 competent cells E. coli Top10⁺ by electroporation method.
 The cDNA libraries constructed by this method are
 full-length enriched cDNA library."

ORIGIN

Alignment Scores:

Pred. No.: 6,33e-94 Length: 582
 Score: 904.00 Matches: 175
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 67.61% Indels: 0
 DB: 12 Gaps: 0

US-10-063-735-128 (1-260) x BM737591 (1-582)

QY 86 SerGlnHisSerGlyValAspIleCysIleAsnAlaGlyLeuAlaArgProAspThr 105
 Db 3 TCTCAGCAGCAGCGGTGAGACATCTGCATCAACATGCTGGCTGGCCCGCCCTGACACC 62
 QY 106 LeuLeuSerGlySerThrSerGlyTrpLysAspMetPheAsnValAsnValLeuAlaLeu 125
 Db 63 CTGCTCTCAGCAGCAGCAGTGGTGGAAAGGACATGTTCAATGTGAACGTGCTGCCCTC 122
 QY 126 SerIleCysThrArgGluAlaTyrGlnSerMetLysGluArgAsnValAspAspGlyHis 145
 Db 123 AGCATCTGCACACGGGAAGCCTACCACTCCATGAAGGAGCGGAATGTGGACGATGGGCAC 182
 QY 146 IleIleAsnIleAsnSerMetSerGlyHisArgValLeuProLeuSerValThrHisPhe 165
 Db 183 ATCATTAACATCAATAGCATGCTGGCCACCGAGTGTTACCCCTGCTGTGACCCACTTC 242
 QY 166 TyrSerAlaThrLysTyrAlaValThrAlaLeuThrGluGlyLeuArgGlnGluLeuArg 185
 Db 243 TATAGTCCCAAGATATGCCGTCACCTGCGCTGACAGAGGACTGTAGGCGAAGAGCTTCGG 302
 QY 186 GluAlaGlnThrHisIleArgAlaThrCysIleSerProGlyValValGluThrGlnPhe 205

Db 303 GAGGCCAGACCCATCCGACCGTGCATCTCCAGGTGGTGAGACACATTC 362
 Qy 206 AlaphelysLeuHisAspLysAspProGluLysAlaAlaThrTyrGluGlnMetLys 225
 Db 363 GCCTTCAAACTCCACGACAGGACCTGAGAAGGCAGCTGCCACCTATGAGCAATGAAG 422
 Qy 226 CysLeuLysProGluAspValAlaGluAlaValIleTyrValLeuSerThrProAlaHis 245
 Db 423 TGCTCTCAAAACCGAGGATGTGGCCGAGGCTGTATCTAGCTCTCAGCACCCCGACAC 482
 Qy 246 lleGlnIleGlyAspIleGlnMetArgProThrGluGlnValThr 260
 Db 483 ATCCAGATCGGAGACATCCAGATGAGCCACGAGAGGTGACC 527

RESULT 31
 BM743569
 LOCUS
 DEFINITION K-EST0016711 S1SNU5s1 Homo sapiens cDNA clone S1SNU5s1-5-C01 5',
 mRNA sequence.

ACCESSION
 VERSION BM743569.1 GI:19064898
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Kim,Y.S.

21C Frontier Korean EST Project 2001
 Unpublished (2002)

TITLE
 JOURNAL
 COMMENT
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Rojun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 5 row: C column: 01
 High quality sequence stop: 550.

FEATURES
 Location/Qualifiers

1..550
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S1SNU5s1-5-C01"
 /sex="F"
 /tissue_type="Ascites"
 /cell_type="Lymphoblast-like"
 /clone_lib="S1SNU5s1"
 /lab_host="Top10F"
 /note="Organ: Stomach; Vector: pCNS; Site 1: EcoRI;
 Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
 bacterial alkaline phosphatase (BAP) and then decapped
 with tabacco acid pyrophosphatase (TAP). The decapped
 intact mRNA was ligated with DNA-RNA linker including EcoR
 I site by treatment of T4 RNA ligase and the first strand
 cDNA was synthesized from oligo dt-selected mRNA by
 priming with dt-tailed vector. The dt-tailed vector was
 adjusted to have about 60nt. The cDNA vector was
 circularized with E. coli DNA ligase after digestion of
 EcoRI which site is also included in vector. An RNA strand
 converted to a DNA strand by Okayama-Berg method. The
 obtained cDNA vectors were used for transformation of
 competent cells E. coli Top10F by electroporation method.
 The cDNA libraries constructed by this method are
 full-length enriched cDNA library. After analyzing and
 sequencing about 2,000 ~ 3,000 colonies in original cDNA
 library, the abundant cDNAs were selected and amplified by
 PCR reaction using vector region primer including T7

promoter as 5' primer and N(dt)14 as 3' primer. The PCR
 products were used as template for synthesis of
 biotinylated single stranded RNA by in vitro transcription
 reaction. The synthesized RNA probes were hybridized with
 antisense single stranded cDNAs prepared from original
 library and incubated with avidin-gel. After removing
 DNA-RNA hybrids by centrifuge, the subtracted cDNA
 libraries were constructed by transformation of the
 remaining DNA into competent cells E. coli Top10F with
 electroporation method."

ORIGIN

Alignment Scores:
 Pred. No.: 1,67e-93 Length: 550
 Score: 900.00 Matches: 179
 Percent Similarity: 99.44% Conservative: 0
 Best Local Similarity: 99.44% Mismatches: 1
 Query Match: 67.31% Indels: 1
 DB: 12 Gaps: 0

US-10-063-735-128 (1-260) x BM743569 (1-550)

Qy 81 PheSerAlaIleArgSerGlnHisSerGlyValAspIleCysIleAsnAsnAlaGlyLeu 100
 Db 3 TTCTCAGCTATCCGTTCTCAGCACAGCGGTGTAGACAC-TGCATCAACAATCTGGCTTG 61
 Qy 101 AlaArgProAspThrLeuLeuSerGlySerThrSerGlyTyrLysAspMetPheAsnVal 120
 Db 62 GCCCGGCGCTGACACCCCTGCTCTCAGCGAGCACCGTGGTTGGAAGACACATGTTCAATGTG 121
 Qy 121 AsnValLeuAlaLeuSerIleCysThrArgGluAlaTyrGlnSerMetLysGluArgAsn 140
 Db 122 AACGTGCTGCCCTCAGCATCTGCACACGGGAAGCTTACCAGTCCATGAAGGAGCGAAT 181
 Qy 141 ValAspAspGlyHisIleIleAsnIleAsnSerMetSerGlyHisArgValLeuProLeu 160
 Db 182 GTGGACGATGGGCACATCATTAACATCAATAGCATGTCTGGCCACCGAGTGTATCCCTCG 241
 Qy 161 SerValThrHisPheTyrSerAlaThrLysTyrAlaValThrAlaLeuThrGluGlyLeu 180
 Db 242 TCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTCTACCTGCGCTGCACAGGAGACTG 301
 Qy 181 ArgGlnGluLeuArgGluAlaGlnThrHisIleArgAlaThrCysIleSerProGlyVal 200
 Db 302 AGCGAAGAGCTTCGGGAGGCCCCAGACCCACATCCGAGCCACGTGCATCTCTCCAGTGTG 361
 Qy 201 ValGluThrGlnPheAlaPheLysLeuHisAspLysAspProGluLysAlaAlaThr 220
 Db 362 GTGGAGACACAATTCGCTTCAAACTCCACGACAGGACCCCTGAGAAGGACGCTGCCACC 421
 Qy 221 TyrGluGlnMetLysCysLeuLysProGluAspValAlaGluAlaValIleTyrValLeu 240
 Db 422 TATGAGCAATAGTGTCTCAACCCGAGGATGTGGCCGAGGCTGTATCTACGTCTTC 481
 Qy 241 SerThrProAlaHisIleGlnIleGlyAspIleGlnMetArgProThrGluGlnValThr 260
 Db 482 AGCACCCCGCCACACATCCAGATTGGAGACATCCAGATGAGGCCACCGAGGAGGTGACC 541

RESULT 32

BI464353

LOCUS

DEFINITION

603204005F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5269728 5',

mRNA sequence.

BI464353

VERSION BI464353.1 GI:15255009

KEYWORDS EST.

SOURCE Homo sapiens

ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 757)

NIH-MGC http://mgi.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

BI464353
 603204005F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5269728 5',
 mRNA sequence.

BI464353

VERSION BI464353.1 GI:15255009

KEYWORDS EST.

SOURCE Homo sapiens

ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 757)

NIH-MGC http://mgi.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

Pred. No.:
Score:

Percent Similarity: 75.63% Conservatives: 42
 Best Local Similarity: 60.57% Mismatches: 42
 Query Match: 66.04% Indels: 26
 DB: 14 Gaps: 1

US-10-063-735-128 (1-260) x CF662388 (1-1143)

QY 6 MetGluArgTTPArgAspArgLeuAlaLeuValThrGlyAlaSerGlyGlyIleGlyAla 25
 DB 53 ATGGATCGCTGGAAGAGTGGTGTCTCTGTCACTGGAGCTTCAGTGGGAATCGGAGCT 112

QY 26 AlaValAlaArgAlaLeuValGlnGlnGlyLeuLysValValGlyCysAlaArgThrVal 45
 DB 113 GCAATCGGAAGTCTCTTGTCCAGCATGCGATGAAGAGTGTATCGGTGGCCAGAAATGTG 172

QY 46 GlyAsnIleGluGluAlaAlaGluCysIysSerAlaGlyTyrProGlyThrLeuIle 65
 DB 173 GAGCAATTCAGAACTGGCAGCAGAATGTGTAGTAGTGGATTTCAGCGGCACTCTGTTC 232

QY 66 ProTyrArgCysAspLeuSerAsnGluAspIleLeuSerMetPheSerAlaIleArg 85
 DB 233 CGGTACAAATGTGATCTGTCTAGAGGACGACATGATATCCATGTCTCTCGGATCAAA 292

QY 86 SerGlnHisSerGlyValAspIleCysIleAsnAsnAlaGlyLeuAlaArgProAspThr 105
 DB 293 GTTCAACATCAGGGCATGACGTGTGCATTAATATGCTGGTTGGCTCTCCAGAGCT 352

QY 106 LeuLeuSerGlySerThrSerGlyTyrIysAspMetPheAsnValAsnValLeuAlaLeu 125
 DB 353 CTACTGAGTGGCAAACTAGTGGCTGGAGACTATGATGGATGTGAATGTAATGGCCGTG 412

QY 126 SerIleCysThrArgGluAlaTyrGlnSerMetLysGluArgAsnValAspGlyHis 145
 DB 413 TCAGTGTGCACCGTGGAGCTTACCAGTCCATGAAAGAAAGAAATTCATGATGTCAT 472

QY 146 IleIleAsnIleAsnSerMetSerGlyHisArgValLeuProLeuSerValThrHisPhe 165
 DB 473 ATCATTAATATCAACAGTATTTGTGGACACCGGTCGTCAACCTGCTGATGCACACTTC 532

QY 166 TyrSerAlaThrIysTyrAlaValThrAlaLeuThrGluGlyLeuArgGlnGluLeuArg 185
 DB 533 TACAGTGGCCAGAAATATGCAGTACGGCTCTCACAGAGGTTTGAGGCAAGAGTTACGA 592

QY 186 GluAlaGlnThrHisIleArgAlaThrCysIleSerProGlyValValGluThrGlnPhe 205
 DB 593 GAGGCCAAAACCCACATACGTGCCATGTATTCTCTGCTGTAGTGAGACAGAAATTT 652

QY 206 AlaPheLysLeuHisAspLysAspProGluLysAlaAlaAlaThrTyrGluGlnMetLys 225
 DB 653 GCCTACCGACTCTTTAGCGNAACCCAGAAATAGCTGTGCTACCTACACAGTGTAAAG 712

QY 225 ----- 225

DB 713 GTATAGACAAATCTGATTCCTTACTATACAAAGGNTTGAAGTGCATCATTTATGTGTT 772

QY 226 -----CysLeuLysProGluAspValAlaGluAlaValIleTyrVal 239
 DB 773 TAAGTGTCAATTTACAGTGCCTTACAGCAGTGTGACATCCGACCTCAGTGGGGTAGTTC 832

QY 240 LeuSerThrProAlaHisIleGlnIleGlyAspIleGlnMetArgProThrGluGln 258
 DB 833 CTGAGTGCTCCTCTCATGTGTCAATTTGGTGACATTCAGATCACCACCCANNTGGAGCAG 889

RESULT 34
 BI546418 714 bp mRNA linear EST 05-SEP-2001
 LOCUS 603198719F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5260037 5',
 DEFINITION mRNA sequence.
 ACCESSION BI546418
 VERSION BI546418.1 GI:154333730
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 714)
 NIH-MGC http://mgs.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshituki and Piero Carninci (RIKEN)
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L14M11655 row: f column: 06
 High quality sequence stop: 704.

FEATURES
 source
 1..714
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5260037"
 /tissue_type="hippocampus"
 /lab_host="DH10B"
 /clone_lib="NIH MGC 95"
 /note="Organ: brain; Vector: pBluescriptR (modified
 pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
 (gtcgag); Oligo-dr primed using primer
 5'-TTTTTTTTTTTNN-3', size-selected for average
 insert size 2.5 kb and normalized to ROT 5. This is a
 primary library enriched for full-length clones and
 constructed using the Cap-trapper method (Carninci, in
 preparation). Library constructed by M. Brownstein
 (NIH/NHGRI, National Institutes of Health). Note: this
 is a NIH_MGC Library."

US-10-063-735-128 (1-260) x BI546418 (1-714)

QY 1 MetAlaArgProGlyMetGluArgTyrArgAspArgLeuAlaLeuValThrGlyAlaSer 20
 DB 129 ATGGCCAGCCCGGCATGGAGCGGTGGCGGCGCGCTG----- 167

QY 21 GlyGlyIleGlyAlaAlaValAlaArgAlaLeuValGlnGlnGlyLeuLysValValGly 40
 DB 168 -----GGCGTGGTCCAGCAGGAGCTGAAGGTGGTGGC 200

QY 41 CysAlaArgThrValGlyAsnIleGluGluLeuAlaAlaGluCysLysSerAlaGlyTyr 60
 DB 201 TGGCCCGGCACTGTGGCGAAC----- 221

QY 61 ProGlyThrLeuIleProTyr-ArgCysAspLeuSerAsnGluGluAspIleLeuSerMe 80
 DB 222 -----ATCGAGATGTGACCTATCAATGAAGAGACATCTCTCCAT 263

QY 80 tPheSerAlaIleArgSerGlnHisSerGlyValAspIleCys-IleAsnAsnAlaGlyL 100
 DB 264 GTTCTCAGCTATCCGTTCTCAGCACACGCGGTGTAGACATCTGCAGTCAACAATGCTGGCT 323

QY 100 euAlaArgProAspThrLeuLeuSerGlySerThrSerGlyTyrLysAspMetPheAsnV 120
 DB 324 TGGCCCGGCGCTGACACCCCTGCTCTCAGGACACCGAGTGGTTGGAAGACATGTTCAATG 383

ORIGIN

Alignment Scores:
 Pred. No.: 3,65e-91 Length: 714
 Score: 881.50 Matches: 187
 Percent Similarity: 81.39% Conservative: 1
 Best Local Similarity: 80.95% Mismatches: 6
 Query Match: 65.93% Indels: 38
 DB: 12 Gaps: 2

QY 120 alAsnValLeuAlaLeuSerIleCysThrArgGluAlaTyrGlnSerMetLysGluArgA 140
 Db 384 TGAACGTGCTGGCCCTCAGCACTCTGCACAGGGAAGCCCTACCAAGCCATGAAGAGCGGA 443
 QY 140 snValAspAspGlyHisIleIleAsnIleAsnSerMetSerGlyHisArgValLeuProL 160
 Db 444 ATGTGGACCATGGCCACATCAATTAACATCAATAGCATGTCTGCCACCCAGGTGTACCCC 503
 QY 160 euSerValThrHisPheTyrSerAlaThrIleTyrValAlaThrAlaLeuThrGluGlyL 180
 Db 504 TGCTGTGTGACCCACTTCTATAGTGCACCAAGATATGCCGTCACTGCGCTGCACAGGGAC 563
 QY 180 euArgGlnGluLeuArgGluAlaGlnThrHisIleArgAlaThrCysIleSerProGlyV 200
 Db 564 TGAGGCAAGAGCTTGGGAGGCCAGACCCACATCCGAGCCAGTGCATCTGCCAGGTG 623
 QY 200 alValGluThrGlnPheAlaPheLysLeuHisAspLysAspProGluLysAlaAlaAlaT 220
 Db 624 TGGTGAGACACAATTCGCTTCAACTCCAGCACAAGACCTGAGAGGCGAGCTGC-A 682
 QY 220 hrTyrGluGlnMetLysCysLysLeuPro 229
 Db 683 CATATGAGCAATGAAGGTGGGCGCTCCC 711

RESULT 35
 AL868596 654 bp mRNA linear EST 03-DEC-2003
 LOCUS AL868596 XGC-egg Silurana tropicalis cDNA clone TBgg124024 5', mRNA
 DEFINITION
 ACCESSION AL868596
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Silurana tropicalis (western clawed frog)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 Xenopodinae; Silurana.
 1 (bases 1 to 654)
 Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
 Sanger Xenopus tropicalis EST project 2001 (11_2003)
 JOURNAL
 COMMENT
 On Sep 15, 2002 this sequence version replaced gi:22888861.
 Contact: Taylor R
 Sanger Institute
 Hinxton, Cambridgeshire, CB10 1SA, UK
 Email: trop@sanger.ac.uk
 Sanger Xenopus tropicalis EST project 2001
 TROPICALIS_SEQUENCE_ID: TEGg124024.pikSP6
 Sequencing primer: SP6
 This sequence is from a Xenopus Gene Collection (XGC) library
 constructed by Aaron M. Zorn.
 cDNA was oligo dT primed from Sug of poly A+ RNA from egg.
 EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the
 5' end and NotI at the 3' end.
 Vector: pCS107; Site 1: EcoRI; Site 2: NotI
 Host: Escherichia coli XLI-blue.
 Location/Qualifiers
 1. .654
 /organism="Silurana tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:9364"
 /clone="TEGg124024"
 /dev_stage="egg"
 /lab_host="Escherichia coli XLI-blue"
 /clone_lib="XGC-egg"
 /note="vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
 was oligo dT primed from Sug of poly A+ RNA from egg.
 EcoRI-NotI cut cDNA was then ligated into pCS107 with
 EcoRI at the 5' end and NotI at the 3' end"

FEATURES
 source
 1001 bp mRNA linear EST 04-AUG-2003
 AGENCOURT_15064616 NICHED_XGC Embs Silurana tropicalis cDNA clone
 IMAGE:6982160 5', mRNA sequence.
 CF225840
 CF225840.1 GI:33426644
 EST.
 Silurana tropicalis (western clawed frog)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 Xenopodinae; Silurana.
 1 (bases 1 to 1001)
 NIH-MGC http://mgs.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892

ORIGIN
 Alignment Scores:
 Pred. No.: 6.09e-91 Length: 654

Score: 879.00 Matches: 168
 Percent Similarity: 89.90% Conservative: 19
 Best Local Similarity: 80.77% Mismatches: 21
 Query Match: 65.74% Indels: 0
 DB: 9 Gaps: 0
 US-10-063-735-128 (1-260) x AL868596 (1-654)
 QY 6 MetGluArgTyrArgAspArgLeuAlaLeuValThrGlyAlaSerGlyGlyIleGlyAla 25
 Db 31 ATGGAGCGCTGGAAGGCGAGGCTGGACCTTGTGACCGGGGCTCGGTGGGCAATCGAGCC 90
 QY 26 AlaValAlaArgAlaLeuValGlnGlnGlyLeuLysValValGlyCysAlaArgThrVal 45
 Db 91 GGGGTTCGCGGGTGTGTTTTCAGCATGGCATGAAAGTGGTGGCTGTCCAGGAGCGTT 150
 QY 46 GlyAsnIleGluGluLeuAlaGluCysLysSerAlaGlyTyrProGlyThrLeuIle 65
 Db 151 GATAAGATTGAAACTGGCTCTGAATGTCAGAGTGTGGCTACCCAGCACCTTATTT 210
 QY 66 ProTyrArgCysAspLeuSerAsnGluGluAspIleLeuSerMetPheSerAlaIleArg 85
 Db 211 CCTTATAAATGTGACCTGTCCAATGAAGAGGAGATTCTGCCATGTTTCACCAATTAAG 270
 QY 86 SerGlnHisSerGlyValAspIleCysIleAsnAlaGlyLeuAlaArgProAspThr 105
 Db 271 ACTTTGTCATCAGGGGTGCGATGTATGTATCAACAATGCAGGCTTGGCCGCGAGCGCT 330
 QY 106 LeuLeuSerGlySerThrSerGlyTyrLysAspMetPheAsnValAsnValLeuAlaLeu 125
 Db 331 TTGCTGAGTGGCAACAGAGGAGTGGAGACATGATTGATGTTTAAATGTTTTCGACTC 390
 QY 126 SerIleCysThrArgGluAlaTyrGlnSerMetLysGluArgAsnValAspAspGlyHis 145
 Db 391 AGTATCTGCACAGAGAGGCTTACCATGTCATGAAGAAAGAAATATCATGATGGCCAT 450
 QY 146 IleIleAsnIleAsnSerMetSerGlyHisArgValLeuProLeuSerValThrHisPhe 165
 Db 451 ATCATAAACATCAACAGCATGAGTGGCCATAGATTCTTCTTCTACAGTTATGCACATT 510
 QY 166 TyrSerAlaThrLysTyrAlaValThrAlaLeuThrGluGlyLeuArgGlnLeuArg 185
 Db 511 TATTTCAGTACTAAGTATGCTGTAACTGCCCTGACAGAGGCGCTCAGGCAAGAGCTCAGA 570
 QY 186 GluAlaGlnThrHisIleArgAlaThrCysIleSerProGlyValValGluThrGlnPhe 205
 Db 571 GAAGAAAGAGTGCATCCGAGCAACGAGTATATCGCCAGGCGCTTGGAAACCTGGATT 630
 QY 206 AlaPheLysLeuHisAspLysAsp 213
 Db 631 GCATTTAAACTCTCTTGATAATGAT 654

RESULT 36
 CF225840
 LOCUS
 DEFINITION
 AGENCOURT_15064616 NICHED_XGC Embs Silurana tropicalis cDNA clone
 IMAGE:6982160 5', mRNA sequence.
 CF225840
 CF225840.1 GI:33426644
 EST.
 Silurana tropicalis (western clawed frog)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 Xenopodinae; Silurana.
 1 (bases 1 to 1001)
 NIH-MGC http://mgs.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgsaps-r@mail.nih.gov
 Tissue Procurement: Robert M. Grainger
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM14644 row: 1 column: 07
 High quality sequence stop: 674.

FEATURES

Location/Qualifiers
 1..1001
 /organism="Silurana tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="IMAGE:6982160"
 /tissue_type="tadpole"
 /dev_stage="embryo, stages 40-45"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NICHD XGC Emb8"
 /note="Vector: pCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV;
 Cloned unidirectionally. Primer: Oligo dt. Average insert
 size 2.1 kb. Constructed by Invitrogen. Note: This is a
 Xenopus Gene Collection (XGC) library."

ORIGIN

Alignment Scores:
 Pred. No.: 1.08e-89 Length: 1001
 Score: 871.00 Matches: 162
 Percent Similarity: 81.86% Conservative: 32
 Best Local Similarity: 68.35% Mismatches: 43
 Query Match: 65.15% Indels: 0
 DB: 14 Gaps: 0

US-10-063-735-128 (1-260) x CF25840 (1-1001)

QY 22 GlylleGlyAlaAlaValAlaArgAlaLeuValGlnGlnGlyLeuValValGlyCys 41
 DB 5 GGCATCGGAGCGCGGTTGCCCGGGTGTTCAGCATGGCATGAAGTGTGGCGCTG 64
 QY 42 AlaAArgThrValGlyAsnIleGluGluLeuAlaAlaGluCysLeuSerAlaGlyTyrPro 61
 DB 65 GCACGAGCGCTGATGAAGATTGGAACACTGGCTGCTGATGTCAGAGTCTGCTACCCA 124
 QY 62 GlyThrLeuIleProTyrArgCysAspLeuSerAsnGluGluAspIleLeuSerMetPhe 81
 DB 125 GGCACCTTATTCCTTATAATGTGACCTGTCCATGAAGAGGAGATCTCTCCATGTTT 184
 QY 82 SerAlaIleArgSerGlnHisSerGlyValAspIleCysIleAsnAlaGlyLeuAla 101
 DB 185 TCAGCAATAAAGACTTTGCATCAGGGGGTGCATGATGATGATCAACATGACGCTTGCC 244
 QY 102 ArgProAspThrLeuLeuSerGlySerThrSerGlyTyrPlyAspMetPheAsnValAsn 121
 DB 245 CGACCGGAGCGCTTGTGATGCGCAACAGAGGGATGAGAACATGATGATGATTAAT 304
 QY 122 ValLeuAlaLeuSerIleCysThrArgGluAlaTyrGlnSerMetIysGluArgAsnVal 141
 DB 305 GTTCTTGCACTCAGTATCTGCACAGAGAGGGCTACAGTCCATGAAGAAAGAAATATC 364
 QY 142 AspAspGlyHisIleIleAsnIleAsnSerMetSerGlyHisArgValLeuProLeuSer 161
 DB 365 GATGATGGCCATATCATTAACATCAACAGTGTCTTGGCCATATCTACCAATGTGCAAAA 424
 QY 162 ValThrHisPheTyrSerAlaThrLysTyrAlaValThrAlaLeuThrGluGlyLeuArg 181
 DB 425 CAGGCTCAGCTTTATTTGTCACCAAGCATACAGTGCAGCGCGCTCACTGAAGCGATAAGA 484
 QY 182 GlnGluLeuArgGluAlaGlnThrHisIleArgAlaThrCysIleSerProGlyValVal 201
 DB 485 CRAAGAGCTGAGAAATTTGAAGAGCCATATTCGTGTTACAGAGCATTTCCCTCGCCCTGTG 544
 QY 202 GluThrGlnPheAlaPheLysLeuHisAspLysAspProGluLysAlaAlaThrTyr 221

Db 545 GAGACAAATTTGCTTACAGATGTTTTGAAATGACCCGTCAATAGCAGCTACGCTGTAC 604
 QY 222 GluGlnMetIysCysLeuLysProGluAspValAlaGluAlaValIleTyrValLeuSer 241
 DB 605 AAATCAATTAAGTGTCTTGTATCTGGTGATATCCGTAATGCTGTTTATATGCTCTGGGT 664
 QY 242 ThrProAlaHisIleGlnIleGlyAspIleGlnMetArgProThrGluGln 258
 DB 665 ACACCACCTCATGTTTCAGGTTTCATGAATGATTGTGAGACCACTGAGCAA 715

RESULT 37

LOCUS BI765897 804 bp mRNA linear EST 25-SEP-2001
 DEFINITION 603046146F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5186459 5',
 mRNA sequence.
 ACCESSION BI765897
 VERSION BI765897
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (Bases 1 to 804)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgsaps-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM11465 row: 1 column: 12
 High quality sequence stop: 748.

FEATURES

Location/Qualifiers
 1..804
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5186459"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_116"
 /note="Organ: pooled colon, kidney, stomach; Vector:
 pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
 source anonymous pool of 3 colons, age 26 yo male, 49 yo
 female, 71 yo male colon, 46 yo male kidney, and pool of 2
 stomachs, 62 yo male and 70 yo female. Library is
 oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.4 kb,
 insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research Genetics tracking code
 023. Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:
 Pred. No.: 2.15e-89 Length: 804
 Score: 867.00 Matches: 193
 Percent Similarity: 93.33% Conservative: 3
 Best Local Similarity: 91.90% Mismatches: 4
 Query Match: 64.85% Indels: 10
 DB: 12 Gaps: 0

US-10-063-735-128 (1-260) x BI765897 (1-804)

QY 1 MetAlaArgProGlyMetGluArgTyrArgAspArgLeuAlaLeuValThrGlyAlaSer 20
 DB 194 ATGCCAGGCGCCGGCATGTGAGCGGTGCGGACCGCGCTGGCGTGTGTGACGGGGCGCTCG 253

```

QY 21 GlyGlyIleGlyAlaAlaValAlaAlaArgAlaLeuValGlnGlnGlyLeuLysValValGly 40
Db 254 GGGGGCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 313
QY 41 CysAlaArgThrValGlyAsnIleGluGluLeuAlaAlaGluCysLysSerAlaGlyTyr 60
Db 314 TGGCCCGCCACGTGGGCAACATCGAGGAGCTGGCTGCTGAATGTAAGAGTGCAGGCTAC 373
QY 61 ProGlyThrLeuLeuProTyrArgCysAspLeuSerAsnGluGluAspIleLeuSerMet 80
Db 374 CCGGGACCTTGCATCCCTACAGATGACCTATCAATGAAGAGACATCCTCTCCATG 433
QY 81 PheSerAlaIleArgSerGlnHisSerGlyValAspIleCysIleAsnAlaGlyLeu 100
Db 434 TTCTCAGCTATCGTTCTCAGCACAGCGGTGTAGACATCTGCATCAACATGCTGGCTTG 493
QY 101 AlaArgProAspThrLeuLeuSerGlySerThrSerGlyTyrPlyAspMetPheAsnVal 120
Db 494 GACCGGCTTGACACCTGCTCTCAGGACACACAGTGGTGGAGGACATGTTCAATGTG 553
QY 121 AsnValLeuAlaLeuSerIleCysThrArgGluAlaTyrGlnSerMetLysGluArgAsn 140
Db 554 AACGTGCTGGCCCTCAGCATCTGCACAGGAGGCTACCATGCCATGAAGGAGCGGAAT 613
QY 141 Val--AspAspGlyHisIleLeuAsnIleAsnSerMetSerGlyHisArgValLeuProL 160
Db 614 GTGGAGACGATGGGCACATCAATACATCAATAGCATGTCTGCCACCGAGTGTACCC 673
QY 160 eu-SerValThrHisPheTyr-SerAla-ThrLysTyrAla-ValThr-AlaLeuThr-G 178
Db 674 TGGTCTGAGCCACTTCTATAAGTCCCAACCAAGTATGCCCGTCACTGGCGCTGACAAG 733
QY 178 LuGlyLeuArgGlnGlu-LeuArgGlu-AlaGlnThrHisIleArgAlaThrCysIleSe 197
Db 734 AGGACTGAGGACAGACTTCGGGAGGCCCAACCAACCAATCGAGACAAAGTGCATCTC 793
QY 197 rProGlyVal 200
Db 794 TCCAGGTGTT 803

RESULT 38
BX773551/c
LOCUS
DEFINITION BX773551 XGC-egg Silurana tropicalis cDNA clone Tegg038a23 3', mRNA
sequence.
ACCESSION BX773551
VERSION BX773551.1 GI:39680756
KEYWORDS EST.
SOURCE
ORGANISM Silurana tropicalis (western clawed frog)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Silurana.
1 (bases 1 to 935)
Croning,W.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
Sanger Xenopus tropicalis EST project 2001 (11_2003)
Unpublished (2003)
Contact: Croning MDR
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: tropesanger.ac.uk
TROPICALIS SEQUENCE ID: Tegg038a23.q1k77
Sequencing primer: T7
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
cDNA was oligo dT primed from Sug of poly A+ RNA from egg.
ECORI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the
5' end and NotI at the 3' end.
Vector: pCS107; Site_1: EcoRI; Site_2: NotI
Host: Escherichia coli XL1-blue.
Location/Qualifiers
1..935

```

```

/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TEgg038a23"
/dev_stage="egg"
/lab_host="Escherichia coli XL1-blue"
/clone_lib="XGC-egg"
/notes="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
was oligo dT primed from Sug of poly A+ RNA from egg.
ECORI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end"

ORIGIN
Alignment Scores:
Pred. No.: 1.99e-87 Length: 935
Score: 851.00 Matches: 158
Percent Similarity: 81.5% Conservative: 32
Best Local Similarity: 67.8% Mismatches: 43
Query Match: 63.6% Indels: 0
DB: 13 Gaps: 0

US-10-063-735-128 (1-260) x BX773551 (1-935)

QY 26 AlaValAlaArgAlaLeuValGlnGlnGlyLeuLysValValGlyCysAlaArgThrVal 45
Db 933 GCGGTTGCCCGGCTGCTTGTTCAGCATGCGATGAAGTGGTGGCTGTGCCAGGAGCGTT 874
QY 46 GlyAsnIleGluGluLeuAlaAlaGluCysLysSerAlaGlyTyrProGlyThrLeuIle 65
Db 873 GATAAGATTGAGAACTGGCTGCTGAATGTCAGATGCTGCTACCCAGGACCTTATT 814
QY 66 ProTyrArgCysAspLeuSerAsnGluGluAspIleLeuSerMetPheSerAlaIleArg 85
Db 813 CCTTATAATGTGACCTGTCCAATGAAGAGGAGATTCTGTCCATGTTTTTCAGATAAAG 754
QY 86 SerGlnHisSerGlyValAspIleCysIleAsnAlaGlyLeuAlaArgProAspThr 105
Db 753 ACTTTGCATCAGGGGGTGGATGTATGATCAACATTCAGGCTTTGGCCCGAGCGCT 694
QY 106 LeuLeuSerGlySerThrSerGlyTyrLysAspMetPheAsnValAsnValLeuAlaLeu 125
Db 693 TTGCTGAGTGGCAAAACAGAGGATGAGAAACAATGATGATGATGATGATGATGATGATG 634
QY 126 SerIleCysThrArgGluAlaTyrGlnSerMetLysGluArgAsnValAspAspGlyHis 145
Db 633 AGTATCTGCACAGAGAGGCTTACCACTCCATGAAGGAAGGATATCGATGATGCGCAT 574
QY 146 IleIleAsnIleAsnSerMetSerGlyHisArgValLeuProLeuSerValThrHisPhe 165
Db 573 ATCAATAACATCAACAGTGTCTTTGGCCATATCTACCAATGTGCAAAACAGGCTCACTTT 514
QY 166 TyrSerAlaThrLysTyrAlaValThrAlaLeuThrGluGlyLeuArgGlnGluLeuArg 185
Db 513 TATTGTGCCACCAAGCATACAGTGCAGCGCTCTCTGACCGGATAGACAGAGCTGAGA 454
QY 186 GluAlaGlnThrHisIleArgAlaThrCysIleSerProGlyValValGluThrGlnPhe 205
Db 453 GAATTGAAGAGCCATATTCGTGTTACAAGCATTTCCCTCGCTCTGTGGAGACAGAAATT 394
QY 206 AlaPheLysLeuHisAspLysAspProGluLysAlaAlaAlaThrTyrGluGlnMetLys 225
Db 393 GCTTACAGATGTTTTGAAAATGACCCGTCATCAATAGCAGCTACGCTGTACAATCAATTAAG 334
QY 226 CysLeuLysProGluAspValAlaAlaGluAlaValIleTyrValLeuSerThrProAlaHis 245
Db 333 TGTCTTGATCCTGGTGATCGCTAATGCTGTTTTATATGCTCTGGGTACACACCTCAT 274
QY 246 IleGlnIleGlyAspIleGlnMetArgProThrGluGln 258
Db 273 GTTCAGGTTTCATGAATGATTGTGAGACCACTGACCAA 235

RESULT 39
BF581815

```

```

LOCUS      BF581815              946 bp      mRNA      linear      EST 12-DEC-2000
DEFINITION 602099816F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4219994 5',
            mRNA sequence.
ACCESSION  BF581815
VERSION     BF581815.1 GI:11655527
KEYWORDS    Mus musculus (house mouse)
SOURCE      Mus musculus
ORGANISM    Mus musculus
REFERENCE   1 (bases 1 to 946)
AUTHORS     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE       Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
JOURNAL     NIH-MGC http://mgc.nci.nih.gov/
COMMENT     National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LAM9802 row: 0 column: 03
            High quality sequence stop: 646.
            Location/Qualifiers
              1..946
                /organism="Mus musculus"
                /mol_type="mRNA"
                /strain="FVB/N"
                /db_xref="taxon:10090"
                /clone="IMAGE:4219994"
                /lab_host="DH10B (T1 phage-resistant)"
                /clone_lib="NCI CGAP Co24"
                /note="Organ: colon; Vector: pCMV-SPORT6; Site 1: NotI;
                Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                Average insert size 1.6 kb. Constructed by Life
                Technologies. Note: this is a NCI_CGAP Library."
FEATURES             source
  source
  RESULT 40
  LOCUS   BI761329
  DEFINITION 603044093F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5184479 5',
            mRNA sequence.
  ACCESSION BI761329
  VERSION    BI761329.1 GI:15752907
  KEYWORDS   Homo sapiens (human)
  SOURCE      Homo sapiens
  ORGANISM    Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  NIH-MGC http://mgc.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  Tissue Procurement: Life Technologies, Inc.
  cDNA Library Preparation: Life Technologies, Inc.
  cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
  DNA Sequencing by: Incyte Genomics, Inc.
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  http://image.llnl.gov
  Plate: LAM11460 row: i column: 24
  High quality sequence stop: 724.
  Location/Qualifiers
    1..748
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:5184479"
      /lab_host="DH10B"
      /clone_lib="NIH MGC 116"
      /note="Organ: pooled colon, kidney, stomach; Vector:
      pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
      source anonymous pool of 3 colons, age 26 yo male, 49 yo
      female, 71 yo male colon; 46 yo male kidney, and pool of 2
      stomachs, 62 yo male and 70 yo female. Library is
      oligo-dT primed and directionally cloned (EcoRV site is
      destroyed upon cloning). Average insert size 1.4 kb,
      insert size range 1-3 kb. Library is normalized and
      enriched for full-length clones and was constructed by C.
  Alignment Scores:
  Pred. No.: 7,71e-87 Length: 946
  Score: 846.00 Matches: 201
  Percent Similarity: 81.37% Conservative: 13
  Best local Similarity: 76.43% Mismatches: 39
  Query Match: 63.28% Indels: 13
  DB: 10 Gaps: 3
  US-10-063-735-128 (1-260) x BF581815 (1-946)
  QY      1 MetAlaArgProGlyMetGluArgTrpArgAspArgLeuAlaLeuValThrGlyAlaSer 20
  Db      66 ATGACTAGAGCTGGCATGGAGCGGTGGCGGACCGGCTGGCACTGGTGACGGGAGCGCTCG 125
  QY      21 GlyGlyIleGlyAlaAlaValAlaArgAlaLeuValGlnGlnGlyLeuLysValValGly 40
  Db      126 GGGGGCATCGGTGGCGCCGTGGCCCGGGCATTAGTCCAGAGGAGTGAAGGTGTGGGT 185
  QY      41 CysAlaArgThrValGlyAsnIleGluGluLeuAlaAlaGluCysLysSerAlaGlyTyr 60
  Db      186 TGTGCCCGCACCGTTGGCAACATCGAGGAGCTGGCTGCTGAATGTAAGATGTCAGGCTAC 245
  QY      61 ProGlyThrLeuIleProTyrArgCysAspLeuSerAsnGluAlaAspIleLeuSerMet 80
  Db      246 CCGGGACTTTGATCCCTACAGATGTGACCTGTCAATGAGGAGGACATCTCTCCATG 305
  QY      81 PheSerAla-IleArgSerGlnHisSerGlyValAlaAspIleCysIleAsnAlaGlyLe 100
  Db      306 TTCTCAGCATGTCGATCCGACACAGTGGCGGTGGATATCTGATCAACATGCCGGCAT 365
  QY      100 uAlaArgProAspThrLeuLeuSerGlySerThrSerGlyTrpLysAspMetPheAsnVa 120
  Db      366 GGCCG--GCTGACACCCCTGCTCTCGGGCAGCACCAGCGGATGGAAGACATGTTCAATGT 423

```

```

QY      120 lAsnValLeuAlaLeuSerIleCysThrArg-GluAlaTyrGlnSerMetLysGluArgA 140
Db      424 GAATGTGCTGCCCTCAGCATCTGCACCTCTGGGAGGCTTATCAGTCCATGAGGAGCGGA 483
QY      140 snValAspAspGlyHisIleIleAsn-IleAsnSerMetSerGlyHisArgValLeuPro 159
Db      484 ACATAGACGAGCGGCACATCATTAACAATCAACAGCATGTGTGGCCACCGAGTCCACCC 543
QY      160 LeuSerValThrHisPhe-TyrSerAlaThrLysTyrAlaValThrAlaLeuThr-GluG 179
Db      544 CAGTCTGTGATCCATTTCTAATAGTGGACTAAGTATGCCGTCCACTGACACTGACAAGAGG 603
QY      179 lLeuArgGlnGluLeu-ArgGluAlaGlnThrHisIleArgAlaThrCysIleSerPro 198
Db      604 GACTCAGGCAAGACTTCTTGGAGGCCAGAGCCATATCCGGCAACGCTGATTCTTCAG 663
QY      199 GlyValValGluThrGlnPheAlaPheLysLeuHisAspLysAspProGluLysAlaAla 218
Db      664 GCTTGGTTAGAGACACACAGTTCGCTTCCACTCATGAACAGACCGGGGAA---GCAGGT 720
QY      219 AlaThrTyrGluGlnMetLysCysLeuLysProGluAspValAlaGluAlaValIleTyr 238
Db      721 GCCCC-TATGAACACGTA---TGTGTTTCAGACAGAGACGCTGCTTGGGCTGTCTATATAG 776
QY      239 ValLeuSerThrProAlaHisIleGlnIleGlyAspIleGlnMetArgProThrGlu 257
Db      777 TTACTTAGGAGACCCCA-----CAGTGGTGGACACTCACATCGCCCGCAGCAG 826

```

```

RESULT 40
LOCUS     BI761329
DEFINITION 603044093F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5184479 5',
            mRNA sequence.
ACCESSION BI761329
VERSION    BI761329.1 GI:15752907
KEYWORDS   Homo sapiens (human)
SOURCE      Homo sapiens
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM11460 row: i column: 24
High quality sequence stop: 724.
Location/Qualifiers
  1..748
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:5184479"
    /lab_host="DH10B"
    /clone_lib="NIH MGC 116"
    /note="Organ: pooled colon, kidney, stomach; Vector:
    pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
    source anonymous pool of 3 colons, age 26 yo male, 49 yo
    female, 71 yo male colon; 46 yo male kidney, and pool of 2
    stomachs, 62 yo male and 70 yo female. Library is
    oligo-dT primed and directionally cloned (EcoRV site is
    destroyed upon cloning). Average insert size 1.4 kb,
    insert size range 1-3 kb. Library is normalized and
    enriched for full-length clones and was constructed by C.

```

FEATURES
source

Contact: Andrew R. Cossins

521 ATTGATGATGCTATATCATATATATCAACAGTATTGTTGGACACCGGTCGTCAACCAT 580
Db
161 SerValThrHisPheTyrSerAlaThrLysTyrAlaValThrAlaLeuThrGluGlyLeu 180
Qy
581 GCTGATGCACACTTCTACAGTGCAGCAATATATGACGTGACGGCTCTCACAGAGGTTTG 640
Db
181 ArgGlnGluLeuArgGluAlaGlnThrHisIleAlaThrCysIleSerProGlyVal 200
Qy
641 AGGCAAGAGTTACGAGAGGCCAAACCCACATACGTGCCATGATATTCTCTCGGCTTA 700
Db
201 ValGluThrGlnPheAlaPheLysLeuHisAspLysAspProGluLysAlaAlaAlaThr 220
Qy
701 GTGGAGACAGAAATTGGCTACCGACTCTTTAGCGAANACCAGGAATAGCTGTTGCTACC 760
Db
221 TyrGluGlnMetLysCysLeuLysProGluAspValAlaGluAlaValIleTyrValLeu 240
Qy
761 TACACAAGTGAAGTGCCTACAGCCATTGACCTCGCCAACTCAGTGGTGTATGCTCTG 820
Db
241 SerThrProAlaHisIleGlnIleGlyAspIleGln 252
Qy
821 AAGCTCCCTCCCTGTTCCCAATTGGGACATTGAGA 856
Db
RESULT 42
AL845993 646 bp mRNA linear EST 20-NOV-2003
LOCUS
DEFINITION
AL845993 XGC-egg Silurana tropicalis cDNA clone TEgg062g24 5', mRNA
sequence.
ACCESSION
AL845993
VERSION
AL845993.2 GI:38487078
KEYWORDS
EST.
ORGANISM
Silurana tropicalis (western clawed frog)
Silurana tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Silurana.
REFERENCE
1 (bases 1 to 646)
Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
Sanger Xenopus tropicalis EST project 2001 (11_2003)
TITLE
Unpublished (2003)
JOURNAL
On Sep 15, 2002 this sequence version replaced gi:22866258.
COMMENT
Contact: Taylor R
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS SEQUENCE ID: TEGg062g24.p1kSP6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn. 5ug of poly A+ RNA from egg.
cDNA was oligo dt primed from 5ug of poly A+ RNA from egg.
EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the
5' end and NotI at the 3' end.
Vector: pCS107; Site 1: EcoRI; Site 2: NotI
Host: Escherichia coli XL1-blue.
Location/Qualifiers
FEATURES
source
1..646
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TEGg062g24"
/dev_stage="egg"
/lab_host="Escherichia coli XL1-blue"
/clone_lib="XGC-egg"
/note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
was oligo dt primed from 5ug of poly A+ RNA from egg.
EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end"

521 ATTGATGATGCTATATCATATATCAACAGTATTGTTGGACACCGGTCGTCAACCAT 580
Query Match: 81.44% Mismatches: 18
DB: 62.23% Indels: 0
Gaps: 0
US-10-063-735-128 (1-260) x AL845993 (1-646)
Qy 6 MetGluArgTrpArgAspArgLeuAlaLeuValThrGlyAlaSerGlyGlyIleGlyAla 25
Db 63 ATGGAGCGCTGGAGCGGCGGCTGTTGACCGGGGCTCGGTGGGCATCGGAGCC 122
Qy 26 AlaValAlaArgAlaLeuValGlnGlnGlyLeuLysValValGlyCysAlaArgThrVal 45
Db 123 GCGGTGCGCGGTGCTGTTGTCAGCATGGCAAGTGGTGGCTGTGCGGAGCGTT 182
Qy 46 GlyAsnIleGluGluLeuAlaAlaGluCysLysSerAlaGlyTyrProGlyThrLeuIle 65
Db 183 GATAAAGATTGAGAACTGGCTGCTGAATGTCAAGTGTGCTGCTGCTGCTGCTGCTTATT 242
Qy 66 ProTyrArgCysAspLeuSerAsnGluGluAspIleLeuSerMetPheSerAlaIleArg 85
Db 243 CCTTATAAATGTGACCTGTCCATGAAGAGAGAGATTCTGTCATGTTTTCAGCAATAAAG 302
Qy 86 SerGlnHisSerGlyValAspIleCysIleAsnAlaGlyLeuAlaArgProAspThr 105
Db 303 ACTTTGTCATCAGGGGCTGATGTATGATCAACAATGACGGCTTGGCCGACCGAGCCT 362
Qy 106 LeuLeuSerGlySerThrSerGlyTyrLysAspMetPheAsnValAsnValLeuAlaLeu 125
Db 363 TTGCTGATGGCAAAACAGAGGGATGGAGAACAAATGATGATGTTTAATGTTTCTTGCATC 422
Qy 126 SerIleCysThrArgGluAlaTyrGlnSerMetLysGluArgAsnValAspAspGlyHis 145
Db 423 AGTATCTGCACAGAGAGCCCTACAGTCCATGAAGAAAGAAATATCGATGATGCCAT 482
Qy 146 IleIleAsnIleAsnSerMetSerGlyHisArgValLeuProLeuSerValThrHisPhe 165
Db 483 ATCATAAACATCAACAGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 542
Qy 166 TyrSerAlaThrLysTyrAlaValThrAlaLeuThrGluGlyLeuArgGlnGluLeuArg 185
Db 543 TATTGAGTACTAAGTATGCTGTAACCTGCCCTGACAGAGGGGCTCAGGCAGAGGCTCAGA 602
Qy 186 GluAlaGlnThrHisIleArgAlaThrCysIleSerProGly 199
Db 603 GAAGAAAGAGTGCATCCGAGCAACGATATATCGCCAGGC 644
RESULT 43
CD469548 672 bp mRNA linear EST 04-JUN-2003
LOCUS
DEFINITION
LeukoS2_4_F10_g1_A024 Stimulated peripheral blood leukocytes S2
Equus caballus cDNA clone LeukoS2_4_F10_A024 5', mRNA sequence.
ACCESSION
CD469548
VERSION
CD469548.1 GI:31390816
KEYWORDS
EST.
SOURCE
Equus caballus (horse)
ORGANISM
Equus caballus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
REFERENCE
1 (bases 1 to 672)
Moore,J.N., Liang,C., Sun,F., Sullivan,R., Shah,M. and Pratt,J.H.
Vandenplas,M., Cordonnier-Pratt,M.-M., Suzuki,Y., Sugano,S.,
An EST database from equine (Equus caballus) stimulated peripheral
blood leukocytes
Unpublished (2003)
Other ESTs: LeukoS2_4_F10.bl_A024
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in

the Human Genome Center, University of Tokyo Institute of Medical Science; tissue and RNA were prepared in the Department of Large Animal Medicine, University of Georgia; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.

Seq primer: Sug5 (CTTCTGCTCAAAAGCTGGC).

FEATURES

source

1. .672
 /organism="Equus caballus"
 /mol_type="mRNA"
 /strain="thoroughbred"
 /db_xref="taxon:9796"
 /clone="LeukoS2_4_F10_A024"
 /sex="male"
 /tissue_type="blood"
 /cell_type="leukocytes"
 /lab_host="DH10B-T1 phage-resistant E. coli"
 /clone_lib="Stimulated peripheral blood leukocytes S2"
 /note="Organ: circulatory system; Vector: pME18S-FL3; site 1: XhoI; site 2: XhoI. The library was prepared from polyA+ RNA from equine peripheral blood leukocytes isolated from a healthy adult horse. The leukocytes were stimulated for 4 hr with 10 ng/ml E. coli O55:B5 LPS. Double-stranded cDNA was cloned unidirectionally into different DraIII sites of the pME18S-FL3 vector (5'-prime DraIII site is CACTGTGTG, 3'-prime DraIII site is CACCATGTG). XhoI excises the cDNA insert."

ORIGIN

Alignment Scores:

Pred. No.: 1.49e-84 Length: 672
 Score: 824.00 Matches: 170
 Percent Similarity: 69.72% Conservative: 5
 Best Local Similarity: 67.73% Mismatches: 6
 Query Match: 61.63% Indels: 70
 DB: 14 Gaps: 1

US-10-063-735-128 (1-260) x CD469548 (1-672)

QY 1 MetAlaArgProGlyMetGluArgTrpArgAspArgLeuAlaLeuValThrGlyAlaSer 20
 Db 130 ATGCCAGGCGCGGATGGAGCGGTGGCTGACCGCTGGCACTGGTAACCGGAGCTCG 189
 QY 21 GlyGlyLeuGlyAlaAlaValAlaAargAlaLeuValGlnGlnGlyLeuLysValValGly 40
 Db 190 GGGGGCATTGGCGGCGCGTACCGCGCCCTGGTCCAGCGGACCTTAAGGTGGTGGGT 249
 QY 41 CysAlaArgThrValGlyAsnIleGluGluLeuAlaAlaGluCysLysSerAlaGlyTyr 60
 Db 250 TGC CGCGGACCGTGGGCAACATCGAG----- 276
 QY 61 ProGlyThrLeuIleProTyrArgCysAspLeuSerAsnGluGluAspIleLeuSerMet 80
 Db 276 ----- 276
 QY 81 PheSerAlaIleArgSerGlnHisSerGlyValAspIleCysIleAsnAsnAlaGlyLeu 100
 Db 276 ----- 276
 QY 101 AlaArgProAspThrLeuLeuSerGlySerThrSerGlyTrpLysAspMetPheAsnVal 120
 Db 277 -----GTG 279
 QY 121 AsnValLeuAlaLeuSerIleCysThrArgGluAlaTyrGlnSerMetLysGluArgAsn 140
 Db 280 AACGTGTCGCCCTCAGCATCTGCACACGGGAGCCTACCATCTCATGAAGGCGGAAA 339
 QY 141 ValAspAspGlyHisIleIleAsnIleAsnSerMetSerGlyHisArgValLeuProLeu 160
 Db 340 GTGATGACGGGCACATCATTAACATCAACAGCATGTCTGCCACCGAGTGTACCCCG 399

QY 161 SerValThrHisPheTyrSerAlaThrLysTyrAlaValThrAlaLeuThrGluGlyLeu 180
 Db 400 TCATGACCACTTCTTATAGTCCCAAGTATGCCGTCACTCACTGACAGAGGACTG 459
 QY 181 AtgGlnGluLeuArgGluAlaGlnThrHisIleAtgAlaThrCysIleSerProGlyVal 200
 Db 460 AGCAAGAGCTTCGGAGGCGCCAGACCCACATCCGAGCCACGTCGCAATTTCTCCAGGATTG 519
 QY 201 ValGluThrGlnPheAlaPheLysLeuHisAspLysAspProGluLysAlaAlaThr 220
 Db 520 GTGGAGACACAGTTCGCTTCAACTCCACGACAGGACCTTGAGAAACAGCTGCCACC 579
 QY 221 TyrGluGlnMetLysCysLeuLysProGluAspValAlaGluAlaValIleTyrValLeu 240
 Db 580 TATGAGCATAAAGTGTCTCAAGCTGAGGATGTGGCGCGGCTGTCATCTATGTCTTC 639
 QY 241 SerThrProAlaHisIleGlnIleGlyAspIle 251
 Db 640 AGCACCGCCCCCAATGTCAGATTGGAGACATC 672

BE869917 614 bp mRNA linear EST 20-OCT-2000
 601446563F1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3850574 5',
 mRNA sequence.
 BE869917
 BE869917.1 GI:10318693
 EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 614)
 NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1AM9569 row: n column: 15
 High quality sequence stop: 602.
 Location/Qualifiers
 1. .614
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3850574"
 /tissue_type="adenocarcinoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_65"
 /notes="Organ: colon; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.8 kb. Library constructed by Life
 Technologies."

FEATURES

source

Alignment Scores:
 Pred. No.: 5.33e-83 Length: 614
 Score: 810.00 Matches: 175
 Percent Similarity: 97.77% Conservative: 0
 Best Local Similarity: 97.77% Mismatches: 4
 Query Match: 60.58% Indels: 4
 DB: 10 Gaps: 0
 US-10-063-735-128 (1-260) x BE869917 (1-614)

ORIGIN

QY 1 MetAlaArgProGlyMetGluArgTrpArgAspArgLeuAlaLeuValThrGlyAlaSer 20

Oligo dT. Average insert size 1.95 kb. Constructed by ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

```

Db      80  ATGCCAGGCGCGCATGGAGCGTGGCGACCGGTGGCGCTGGGTGACGGGGGCGCTCG 139
Qy      21  GlyGlyIleGlyAlaAlaValAlaArgAlaLeuValGlnGlnGlyLeuLysValValGly 40
Db      140  GGGGGCATCGCGCGCGCGTGGC-CGGGCTTGGTCCAGCAGGAGACTGAAGTGTGGGC 198
Qy      41  CysAlaArgThrValGlyAsnIleGluLeuAlaAlaGluCysLysSerAlaGlyTyr 60
Db      199  TGGCGCGCATCTGTGGGCAACATCAGGAGCTGGCTGTGAATGTAAAGTGCAGGCTAC 258
Qy      61  ProGlyThrLeuIleProTyrArgCysAspLeuSerAsnGluAlaGluAspIleLeuSerMet 80
Db      259  CCGGGGATTTGATCCCTTACAGATGTGACCTATCAATGAAGAGGACATCTCTCCATG 318
Qy      81  PheSerAlaIleArgSerGlnHisSerGlyValAspIleCysIleAsnAsnAlaGlyLeu 100
Db      319  TTCTCAGTATCCGTTCTCAGCACAGCGGTGTAGACATCTGCATCAACAATGCTGGCTTG 378
Qy      101  AlaArgProAspThrLeuLeuSerGlySerThrSerGlyTyrTrpLysAspMetPheAsnVal 120
Db      379  GCGCGGCTCAGACCCCTCTCTCAGGACAGCACCAGTGGTGTGAAGGACATGTTCATGTG 438
Qy      121  AsnValIleuAlaLeuSerIleCysThrArgGluAlaTyrGlnSerMetLysGluArgAsn 140
Db      439  AACGTGCTGGCCCTCAGCATCTGCACACGGGAGGCTACCACTGTCATGAAGAGCGGAAT 498
Qy      141  ValAspAspGlyHisIleIleAsnIleAsnSerMetSerGlyHisArgValLeuProLeu 160
Db      499  GTGGACGATGGGCATCATTAACATCAATAGCATGTCTGGCCACGA-GTGTACCCCTG 557
Qy      161  SerValThrHisPheTyrSerAlaThrLysTyrAlaValThrAlaLeuThrGluGly 179
Db      558  TCTGTGACGCATTC-TATAGTGCCACCAAGTAGTC-GTCACCTGCGCTGACAGAGGGA 612

RESULT 45
BU839980
LOCUS      BU839980
DEFINITION AGENCOURT_8947516 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6329956
5', mRNA sequence.
ACCESSION BU839980.1 GI:24024375
VERSION    BU839980
KEYWORDS   EST.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
REFERENCE  1 (bases 1 to 977)
AUTHORS   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE     Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
JOURNAL   NIH-MGC http://mgc.nci.nih.gov/
COMMENT   National Institutes of Health, Mammalian Gene Collection (MGC)
          Contact: Robert Strausberg, Ph.D.
          Email: cgaabs-r@mail.nih.gov
          Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman,
          Ph.D.
          cDNA Library Preparation: ResGen, Invitrogen Corp
          CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Agencourt Bioscience Corporation
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: L1AM13783 row: j column: 05
          High quality sequence stop: 612.
          Location/Qualifiers
            1..977
              /organism="Mus musculus"
              /mol_type="mRNA"
              /db_xref="taxon:10090"
              /clone="IMAGE:6329956"
              /lab_host="DH10B (phage-resistant)"
              /clone_lib="NIH_MGC_130"
              /note="Organ: otocysts; Vector: pCMV-SPORT6.1; Site:1;
              EcoRV; Site_2: NotI; Cloned unidirectionally. Primer:

```

ORIGIN

```

Alignment Scores:      2.34e-82      Length:      977
Pred. No.:            807.50      Matches:     164
Score:                85.64%      Conservative: 3
Percent Similarity:    84.10%      Mismatches:  8
Best Local Similarity: 60.40%      Indels:      20
Query Match:          13          Gaps:        1
DB:

US-10-063-735-128 (1-260) x BU839980 (1-977)

Qy      1  MetAlaArgProGlyMetGluArgTyrArgAspArgLeuAlaLeuValThrGlyAlaSer 20
Db      104  ATGACTAGAGTGGCAGCGGTGGCGCCGACCGCTGGCCTGGTGCACGGGAGCCTCG 163
Qy      21  GlyGlyIleGlyAlaAlaValAlaArgAlaLeuValGlnGlnGlyLeuLysValValGly 40
Db      164  GGGGGCATCTGTGGCGCGTGGCCCGGCATTAGTCCAGCAGGAGACTGAAGTGTGGGT 223
Qy      41  CysAlaArgThrValGlyAsnIleGluLeuAlaAlaGluCysLysSerAlaGlyTyr 60
Db      224  TGTGCCGACCGCTTGGCAAC----- 244
Qy      61  ProGlyThrLeuIleProTyr-ArgCysAspLeuSerAsnGluAlaGluAspIleLeuSerMe 80
Db      245  -----ATCGAGATGTGACCTGTCAATGAGGAGGACATCTCTCCAT 286
Qy      80  tPheSerAlaIleArgSerGlnHisSerGlyValAspIleCysIleAsnAsnAlaGlyLe 100
Db      287  GTTCTCAGCTGTCCGATCCAGCACAGTGGGTGGATATCTGCATCAACAATGCGGCGAT 346
Qy      100  uAlaArgProAspThrLeuLeuSerGlySerThrSerGlyTyrTrpLysAspMetPheAsnVa 120
Db      347  GSCCGGCGCTGACACCCCTGCTCTCGGGCAGCACCGCGGATGGAAGGACATGTTCAATGT 406
Qy      120  IasnValIleuAlaLeuSerIleCysThrArgGluAlaTyrGlnSerMetLysGluArgAs 140
Db      407  GAATGTGTGGCCCTCAGCATCTGCACTCGGGAGGCTTATCTCATGTAGAGGAGCGGA 466
Qy      140  nValAspAspGlyHisIleIleAsnIleAsnSerMetSerGlyHisArgValLeuProLe 160
Db      467  CATAGCAGCGCGGCACATCATTAACATCAACAGCATGTGTGGCCACCGAGTCCACCCCA 526
Qy      160  uSerValThrHisPheTyrSerAlaThrLysTyrAlaValThrAlaLeuThrGluGlyLe 180
Db      527  GTCTGTGATCCATTTCTATAGTGGCAGTAAAGTATGCCGTCATGCGCTGACACTGACAGGGGACT 586
Qy      180  uArgGlnGluLeuArgGluAlaGlnThrHisIleArgAlaThr 194
Db      587  CAGGCAAGAGCTTCTGGAGGCCGACGCCATATCCGGGGCCAG 629

RESULT 46
W58472
LOCUS      W58472
DEFINITION IMAGE:341651 5' similar to SW:YOXD_BACSU P14802 HYPOTHETICAL 25.3
KD PROTEIN IN RTP 5'REGION ; mRNA sequence.
ACCESSION W58472
VERSION    W58472.1 GI:1365185
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 567)
REFERENCE  Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
AUTHORS   Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Matra, M.,
          Parsons, J., Rifkin, J., Rohlfing, T., Soares, M., Tan, F.,
          Trevisan, E., Waterston, R., Williamson, A., Wohlmann, P. and
          Wilson, R.

```

TITLE JOURNAL COMMENT

The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1354 Std Error: 0.00
Seq primer: mob.REGA-RT
High quality sequence stop: 351.

FEATURES

Location/Qualifiers
1..567
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:1267026"
/db_xref="taxon:9606"
/clone="IMAGE:341651"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal heart_NBHL19W"
/note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTCACCAATCTGAATGGAGCGCGGCATCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by
M.Fatima Bonaldo. This library was constructed from the
same fetus as the fetal lung library, Soares fetal lung
NBHL19W."

ORIGIN

Alignment Scores:
Pred. No.: 5.11e-82 Length: 567
Score: 801.00 Matches: 175
Percent Similarity: 93.58% Conservative: 0
Best Local Similarity: 93.58% Mismatches: 4
Query Match: 99.91% Indels: 8
DB: 14 Gaps: 0

US-10-063-735-128 (1-260) x W58472 (1-567)

QY 50 GluLeuAlaGluCysLysSerAlaGlyTyrProGlyThrLeuLeuProTyrArgCys 69
|||||
Db 18 GAGTGGCTGCTGAATGTAGAGTGCAGGCTACCCCGGAGCTTTGATCCCTACAGATGT 77
QY 70 AspLeuSerAsnGluGluAspIleLeuSerMetPheSerAlaIleArgSerGlnHisSer 89
|||||
Db 78 GACCTATCAATGAGAGGACATCTCTCCATGTCTCAGCTATCCGCTCTCAGCACAGC 137
QY 90 GlyValAspIleCysIleAsnAsnAlaGlyLeuAlaArgProAspThrLeuLeuSerGly 109
|||||
Db 138 GCTGTAGACATCGCATCAACATGCTGGCTTGGCCGGCTGACACCTCTCTCAGGC 197
QY 110 SerThrSerGlyTrpIysAspMetPheAsnValAsnValLeuAlaLeuSerIleCysThr 129
|||||
Db 198 AGCACCAGTGGTGGAGGACATGTTCAATGTGAACGTGTGGCCCTCAGCATCTGCACA 257
QY 130 ArgGluAlaTyrGln-SerMetLysGlu-ArgAsnValAsp-AspGlyHisIle-IleAs 148
|||||
Db 258 CGGGAAGCTTACCAGTCCATGAGGAGCGGAATGTGGACCNATGGGCACATCAATTAA 317
QY 148 nIleAsn-SerMetSerGlyHis-ArgValLeuProLeuSerValThrHisPheTyrSer 167
|||||
Db 318 CATCAATAAGCATGTCTGGCCACCCGAGTGTACCCCTGTCTGTGACCCACTTCTATAGT 377
QY 168 AlaThrLysTyrAlaValThrAlaLeuThrGluLeuArgGlnGluLeuArgGluAla 187

Db 378 GCCACCAAGTATCCGCTCACTGCGCTGACAGAGGACTGAGGCAAGAGCTTCGGAGGCC 437
QY 188 GlnThrHisIleArgAlaThrCysIleSerProGlyVal-ValGluThrGlnPheAlaPh 207
|||||
Db 438 CAGACCCACATCCGAGCCAGCTGCACTCTCCAGGTGTGGTGGAGACACAATTCGCCTT 497
QY 207 eLysLeu-HisAspLysAspProGlyLysAlaAlaAlaThrTyrGluGlnMetLysCysL 227
|||||
Db 498 CAAACTTCCAGCAACAAGGACCCCTGAGANGAGCTGCACCTATGAGCNAATGAGTGTC 557
QY 227 eLys 228
Db 558 TCAA 562

RESULT 47
CF348186
LOCUS
DEFINITION 763 bp mRNA linear EST 18-AUG-2003
AGENCOURT_15225322 NIH_ZGC_10 Danio rerio cDNA clone IMAGE:7001094
5', mRNA sequence.
CF348186
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Danio rerio (zebrafish)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 763)
NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS
TITLE
JOURNAL
COMMENT
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgepbs-remail.nih.gov
Tissue Procurement: Len Zon, Harvard
cDNA Library Preparation: Open Biosystems
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14693 row: n column: 05
High quality sequence stop: 510.

FEATURES

Location/Qualifiers
1..763
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:7001094"
/tissue_type="whole body"
/lab_host="DH10B"
/clone_lib="NIH_ZGC_10"
/notes="Vector: pExpress1; Site 1: NotI; Site 2: EcoRV;
Bulk tissue was collected from a whole adult individual
from the Tuebingen strain. 1st strand cDNA was primed with
a Not I - oligo(dT) primer, double-stranded cDNA was
cloned into the Not I and EcoRV sites of pExpress-1.
Library was size-selected for >1 kb fragments. A
normalized version of this library is also available
(NIH_ZGC_7). Library was constructed by Open Biosystems
(Huntsville, AL)."

ORIGIN

Alignment Scores:
Pred. No.: 2.85e-81 Length: 763
Score: 796.50 Matches: 158
Percent Similarity: 84.96% Conservative: 34
Best Local Similarity: 69.91% Mismatches: 33
Query Match: 59.57% Indels: 3
DB: 14 Gaps: 0

US-10-063-735-128 (1-260) x CF348186 (1-763)

Qy	6	MetGluArgTrpArgAspArgLeuAlaLeuValThrGlyAlaSerGlyGlyIleGlyAla	25
Db	28	TTGGATCGCTGGAAAGCAGAGTGTCTTTGTCTAGCTGGAGCTTCAGTAGGAATAGGAGCT	87
Qy	26	AlaValAlaArgAlaLeuValGlnGlnGlyLeuIysValValGlyCysAlaArgThrVal	45
Db	88	GCAATCCGAAAAAGCTCTTTGTCACAGATGGCATGAAGGTGGTCCGATGTGCCAGAAATGTG	147
Qy	46	GlyAsnIleGluGluLeuAlaAlaGluCysIysSerAlaGlyTyrProGlyThrIleuIle	65
Db	148	GAGCAAAATAGAGAAACTGGCGCTCAATGTCTCAGTGGCGGACTCAGCGGTGTCTGTGTT	207
Qy	66	ProTyrArgCysAspLeuSerAsnGluAlaAspIleLeuSerMetPheSerAlaIleArg	85
Db	208	CCATATAAATGTGATCTTTACGTAGAGATGAAGTTTGTCCAATGTTCTCCGTGATTAAAG	267
Qy	86	SerGlnHisSerGlyValAspIleCysIleAsnAsnAlaGlyLeuAlaArgProAspThr	105
Db	268	GCTCAACATAAGGGTGTGTGATGTGTGCATTAAATATGCTGTTTACGCTCTGCCAGAGCCT	327
Qy	106	LeuLeuSerGlySerThrSerGlyTrpIysAspMetPheAsnValAsnValLeuAlaLeu	125
Db	328	CTGTTGAACCGGCAAAAGCCAGCGGCTGGAGGACTATGATGAACGTGAATGTAAATGGCCGTG	387
Qy	126	SerIleCysThrArgGluAlaIleGlnSerMetIysGluArgAsnValAspAspGlyHis	145
Db	388	GCAGTGTGCACCGGTGAGGCTTTACCACTCCATGAAGAAAGAAATATTGATGATGGCCAT	447
Qy	146	IleIleAsnIleAsnSerMetSerGlyHisArgValLeuProLeuSerValThrHisPhe	165
Db	448	ATCATTAATATATAACAGATATGCTGGGCATCGGGTTGTAAACAGTGCCTATACACATTC	507
Qy	166	TyrSerAlaThrLysTyrAlaValThrAlaLeuThrGluGlyLeuArgGlnGluLeuArg	185
Db	508	TACACCGGTACTAATACGACGTGCTCTCACCGAAGGTTTGAGGCAAGAGTATACGA	567
Qy	186	GluAlaGlnThrHisIleArgAlaThrCysIleSerProGlyValValGluThrGlnPhe	205
Db	568	GAGGCCAAACCCACATACGTGCCAAGTATATCCCTCTGGTTTAAATGGAGACAGAATT	627
Qy	205	ealaphelysLeuHisAspIysAspProGluIysAlaAlaIleThrTyrGluGlnMetIy	225
Db	628	TGCCCTACAAACTCTTTAGTGAANA-CCAAACAAGGCTTCTGCCACCTACGAAGTAA-AA	684
Qy	225	sCysLeuIysProGlu	230
Db	685	GTGCTCTGCAACCAAG	700

RESULT 48	AL8981132	648 bp	mRNA	linear	EST 04-DEC-2003
LOCUS	AL8981132				
DEFINITION	XGC-egg <i>Silurana tropicalis</i> cDNA clone TE99038a23 5', mRNA sequence.				
ACCESSION	AL8981132				
VERSION	AL8981132.2	GI:38697669			
KEYWORDS	EST.				
SOURCE	<i>Silurana tropicalis</i> (western clawed frog)				
ORGANISM	<i>Silurana tropicalis</i>				
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Xenopodinae; <i>Silurana</i> .				
REFERENCE	1 (bases 1 to 648)				
AUTHORS	Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.				
TITLE	Sanger Xenopus tropicalis EST project 2001 (11_2003)				
JOURNAL	Unpublished (2003)				
COMMENT	On Sep 16, 2002 this sequence version replaced gi:22950524.				
	Contact: Taylor R				
	Sanger Institute				
	Hinxton, Cambridgeshire, CB10 1SA, UK				
	Email: trop@sanger.ac.uk				

Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: TEG9038a23.p1kSP6
TROPICALIS_SEQUENCE_ID: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
cDNA was oligo dT primed from 5ug of poly A+ RNA from egg.
EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the
5' end and NotI at the 3' end.
Vector: pCS107; Site 1: EcoRI; Site 2: NotI
Host: Escherichia coli XL1-blue.

FEATURES source

```

1. 648
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db xref="taxon:8364"
/clone="TEGG038a23"
/dev_stage="egg"
/lab_host="Escherichia coli XL1-blue"
/clone_lib="Xc-egg"
/note=vector: PCS107; Site1: ECoRI;
was oligo dT primed from 5' end of poly
ECoRI-NotI cut cDNA was then ligated
5' end of the 5' end and NotI at the 3'

```

ORIGIN

Alignment Scores:		7.12e-81	Length: 648
Pred. No.:	792.00	Matches: 150	
Score:	85.64%	Conservative: 23	
Percent Similarity:	74.26%	Mismatches: 29	
Best Local Similarity:	59.24%	Indels: 0	
Query Match:	9	Gaps: 0	
DB:			
US-10-063-735-128 (1-260) x AL898132 (1-648)			
QY	6	MetGluArgTrpArgAspArgLeuAlaLeuValThrGlyAlaSerGlyGlyIleGlyAla	25
DB	41	ATGAGCGCTGAAAGGCGAGGCTGGCACTTGTGACCGGGGCTCGGTGGGCATCGAGCC	100
QY	26	AlaValAlaArgAlaLeuValGlnGlnGlyLeuLysValValGlyCysAlaArgThrVal	45
DB	101	CGGTTGGCCGGGTCTTGTTCAGCATGGCAATGAAGTGGTGGCTGTGCCAGAGCGTT	160
QY	46	GlyAsnIleGluGluLeuAlaAlaGluCysLysSerAlaGlyTyrProGlyThrLeuIle	65
DB	161	GATAGATTGAAGAACTGGCTGCTGAATGTCAGAGTGTGCTACCCAGGACCTTATT	220
QY	66	ProTyrArgCysAspLeuSerAsnGluAspIleLeuSerMetPheSerAlaIleArg	85
DB	221	CCTTATAATGTGACCTCTCCAATGAGAGGAGATTCTGTCCATGTTTTTCAGCAATAAG	280
QY	86	SerGlnHisSerGlyValAspIleCysIleAsnAsnAlaGlyLeuAlaArgProAspThr	105
DB	281	ACTTTTGCATCAGGGGGTCCGATGTATGTATCAACATGAGGCTTGGCCGACCGAGCCT	340
QY	106	LeuLeuSerGlySerThrSerGlyTryptLysAspMetPheAsnValAsnValLeuAlaLeu	125
DB	341	TTGCTGAGTGGCAAAACAGAGGATCGAACAATGATTGATGTTAATGTTCTTGCACCTC	400
QY	126	SerIleCysThrArgGluAlaTyrGlnSerMetLysGluArgAsnValAspAspGlyHis	145
DB	401	AGTATCTGCACAAGAGAGGCCCTACCAGTCCATGAAGGAAAGGAATATCGATGATGGCCAT	460
QY	146	IleIleAsnIleAsnSerMetSerGlyHisArgValLeuProLeuSerValThrHisPhe	165
DB	461	ATCATAAACATCAACAGTGTCTTCGGCCCATATCTACCAATGTGCAAAAACAGGCTCATT	520
QY	166	TyrSerAlaThrLysTyrAlaValThrAlaLeuThrGluGlyLeuArgGlnGluLeuArg	185
DB	521	TATTTGGCCACCACGATCACATGACGCGGCTCACTGAGCGGATGAAGCAAGAGCTGAGA	580
QY	186	GluAlaGlnThrHisIleArgAlaThrCysIleSerProGlyValValGluThrGlnPhe	205

```

Db      581 GAATTGAGACCATATTGCTGTTTACAAGCAATTCCTCGCTTGTGGAGACAGAATTT 640
QY      206 AlaPhe 207
Db      641 GCTTAC 646

RESULT 49
CD599740
LOCUS   CD599740
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES
source
1..650
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="RK119A4F07"
/dev_stage="mature"
/clone_lib="Zebrafish Kidney Marrow cDNA library"
/notes="Organ: kidney; Vector: pBS-CMV; Site 1: XhoI; Site 2: EcoRI; Total RNA was extracted from the kidney tissues of mature zebrafish. The poly (A)+ RNA fraction was separated from total RNA by oligo (dT) cellulose chromatography. Library was initially constructed in the lambdaZAP Express vector (Stratagene) and in vivo excised into pBS-CMV vector."

ORIGIN
Alignment Scores:
Pred. No.:      3..54e-80      Length:      650
Score:          786.00      Matches:     147
Percent Similarity: 87.37%      Conservative: 26
Best Local Similarity: 74.24%      Mismatches:  25
Query Match:    58.79%      Indels:      0
DB:             14          Gaps:         0

US-10-063-735-128 (1-260) x CD599740 (1-650)

QY      56 MetGluArgTrpArgAspArgLeuAlaLeuValThrGlyAlaSerGlyGlyIleGlyAla 25
Db      56 ATGGATCGCTGGAAAGCAGCAGAGTGTCTTGTCTCACTGGAGCTTCATAGATAGAGCT 115
QY      26 AlaValAlaArgAlaLeuValGlnGlnGlnGlyLeuLysValValGlyCysAlaArgThrVal 45
Db      116 GCATCGCAAAAGCTCTTGTCCAGCATGGCATGAGGTGGTGGATGCCAGAAATGG 175
QY      46 GlyAsnIleGluGluLeuAlaAlaGluCysLysSerAlaGlyTyrProGlyThrLeuIle 65
Db      176 GAGCAATAGAGAACTGGCGCTGAATGTGTGTCAGTGGCGGATACAGCGGTGCTCTGTTT 235

```

```

QY      66 ProTyrArgCysAspLeuSerAsnGluAspIleLeuSerMetPheSerAlaIleArg 85
Db      236 CCATATAAATGTGATCTTTTCAGTAGAGGATGAAGTTTGTCCATGTTCTCTCGATTAG 295
QY      86 SerGlnHisSerGlyValAspIleCysIleAsnAsnAlaGlyLeuAlaArgProAspThr 105
Db      296 GCTCAACATAAGGGTGTGATGTGTGATTAATATATGCTGTTAGCTCTGCCAGAGCCT 355
QY      106 LeuLeuSerGlySerThrSerGlyTyrLysAspMetPheAsnValAsnValLeuAlaLeu 125
Db      356 CTGTTGAACGGCAACAGCCAGCGCTGGAGCACTATGATGAATGAATTTGGCCCTG 415
QY      126 SerIleCysThrArgGluAlaTyrGlnSerMetLysGluArgAsnValAspAspGlyHis 145
Db      416 TCATTGTGCACCGGTGAGGCTTTCAGTCCATGAAAGAAAGAAATATTGATGGCCAT 475
QY      146 IleIleAsnIleAsnSerMetSerGlyHisArgValLeuProLeuSerValThrHisPhe 165
Db      476 ATCATTAATTAACAGTATGTCGGCATCGGGTTGTAACAGTGCCTATACACATTC 535
QY      166 TyrSerAlaThrLysTyrAlaValThrAlaLeuThrGluGlyLeuArgGlnGluLeuArg 185
Db      536 TACACCGCTACTAATACGAGTGACTGCTCTACCCGAGGTTTGAGGCAAGATTACGA 595
QY      186 GluAlaGlnThrHisIleArgAlaThrCysIleSerProGlyValValGluThr 203
Db      596 GAGGCCAAACCCACATACGTCGACAGTGCACCAAGTATATCCCTCGTTTGTGGAGACA 649

RESULT 50
CD474613
LOCUS   CD474613
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES
source
1..863
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:6796850"
/lab_host="DH10B (T1-resistant)"
/clone_lib="NCI CGAP ZKId1"
/notes="Organ: kidney; Vector: pCMV-SPORT6.1; Site 1: EcoRV; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.8 Kb. Constructed by J. Wang (Research Genetics, Invitrogen Corp) from tissue donated by L. Zon (Harvard University). Note: this is a NCI CGAP Library."

ORIGIN
Alignment Scores:

```


Fri Sep 17 09:20:50 2004

Pred. No.: 7.6e-80 Length: 863
 Score: 785.00 Matches: 144
 Percent Similarity: 86.83% Conservative: 34
 Best Local Similarity: 70.24% Mismatches: 26
 Query Match: 58.71% Indels: 1
 DB: 14 Gaps: 0

US-10-063-735-128 (1-260) x CA474613 (1-863)

QY	57	SerAlaGlyTyrProGlyThrLeuIleProTyrArgCysAspLeuSerAsnGluGluasp	76
Db	1	AGTGGAGGACTCAGCGGTGCTCTTTCCATATAAATGTGACCTTCAGTAGAGGATGAA	60
QY	77	IleLeuSerMetPheSerAlaIleArgSerGlnHisSerGlyValAspIleCysIleAsn	96
Db	61	GTTTTGTCATGTTCTCTCGATTAGGCTCAACATAAGGGTGTGATGTGTGCTGCTTAAT	120
QY	97	AsnAlaGlyLeuAlaArgProAspThrLeuLeuSerGlySerThrSerGlyTyrIleAsp	116
Db	121	AATGCTGGTTAGCTCTGCCAGAGCCTCTGTGAACCGCAAAACGCGGTGGAGGACT	180
QY	117	MetPheAsnValAsnValLeuAlaLeuSerIleCysThrArgGluAlaTyrGlnSerMet	136
Db	181	ATGATGAACGTGAATGTAATTTGGCTGGCAGTGTGCACCCGTGAGGCTTACCAGTCCATG	240
QY	137	LysGluArgAsnValAspAspGlyHisIleIleAsnIleAsnSerMetSerGlyHisArg	156
Db	241	AAAGAAAGAAATATTGATGATGGCCATATCATTAATATTAAACAGTATGCTGGGCATCGG	300
QY	157	ValLeu-ProLeuSerValThrHisPheTyrSerAlaThrIleValAlaValThrAlaLe	176
Db	301	GTTGTAAACAGTGCCTTTTACACACTTCTACACCGCTACTAAATACGCGAGTACTGCTCT	360
QY	176	uThrGluGlyLeuArgGlnGluLeuArgGluAlaGlnThrHisIleArgAlaThrCysIle	196
Db	361	CACCGAAGGTTTGAGGCAAGATTACGAGAGGCCAAACCCACATACGTGCCACAAGTAT	420
QY	196	eSerProGlyValValGluThrGlnPheAlaPheLysLeuHisAspLysAspProGluIly	216
Db	421	ATCCCTCTGTTTGTGGAGACAGAAATTTGCCCTACAGACTTTTGTGTAACCAAGACAA	480
QY	216	sAlaAlaAlaThrTyrGluGlnMetLysCysLeuLysProGluAspValAlaGluAlaVa	236
Db	481	GGCTTCTGCCACCTACAAAGATATAAAGTGCTGAGCCAGAGATCTAGCAATGCACT	540
QY	236	IleTyrValLeuSerThrProAlaHisIleGlnIleGlyAspIleGlnMetArgProTh	256
Db	541	GGTTTATGTCCTTAGCGCTCTCTCTCATGTTCAATTTGCTGACATTCAGATGAGACCTGT	600
QY	256	rGluGlnValThr	260
Db	601	GGAACAGCTGACA	613

Search completed: September 16, 2004, 18:36:46
 Job time : 2478 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2004, 10:41:25 ; Search time 1656 Seconds
(without alignments)
12292.829 Million cell updates/sec

Title: US-10-063-735-127

Perfect score: 1505

Sequence: 1 cgcgcatcggaacccaagcag.....tcgcaacacaaaaaa 1505

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database :

EST:

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	772.4	51.3	822	12	BG741589 602635416
2	763.2	50.7	829	12	BG324477 60242421
3	732.4	48.7	1010	12	BM471232 AGENCOURT
4	718.8	47.8	1071	13	BQ278156

5	701.4	46.6	940	12	BI818880
6	698.8	46.4	1063	10	BE796469
7	659	43.8	679	10	BF979853
8	656.4	43.6	672	13	BU624717
9	650	43.2	852	12	BG282361
10	644.8	42.8	980	9	AL568188
11	639.4	42.5	669	14	CB851997
12	638.4	42.4	640	12	BM755987
13	633	42.1	649	12	BM717452
14	627	41.7	627	12	BM848012
15	623.6	41.4	650	12	BM674957
16	621	41.3	787	10	BF967768
17	620.8	41.2	651	9	AV654573
18	618.8	41.1	658	12	BM723373
19	617.8	41.0	886	10	BE895089
20	615.6	40.9	757	12	BI464353
21	615.6	40.9	932	10	BG030249
22	603.4	40.1	923	10	BE617298
23	602	40.0	602	12	BM764342
24	600.8	39.9	804	12	BI765897
25	596.6	39.6	937	13	BU539642
26	586.4	39.0	691	12	BM793929
27	586	38.9	1022	9	AL559036
28	580.4	38.6	582	12	BM737591
29	578.4	38.4	769	10	BE386240
30	573	38.1	573	12	BM742038
31	567.4	37.7	614	10	BE869917
32	566	37.6	566	12	BM744024
33	565	37.5	572	12	BM682228
34	550.2	36.6	748	12	BI761329
35	544.2	36.2	914	12	BG967186
36	543	36.1	570	12	BM675954
37	541	35.9	541	12	BG231973
38	539	35.8	550	12	BM743569
39	533	35.4	566	14	CA424831
40	532.6	35.4	612	13	EX671560
41	532.6	35.4	751	10	BG037061
42	530.2	35.2	714	12	BI546418
43	528	35.1	540	12	BM848396
44	523.2	34.8	528	9	AW001572
45	512	34.0	513	10	BE302814
46	509.8	33.9	546	12	BI260323
47	506.6	33.7	716	10	BE962188
48	504.2	33.5	976	9	AL559035
49	503.6	33.5	881	10	BF233600
50	495.2	32.9	536	14	W58459
51	489	32.5	489	9	AI125225
52	487.2	32.4	550	10	BE931725
53	485.2	32.2	535	14	W58347
54	479.4	31.9	578	10	AW974208
55	478	31.8	516	12	BM708058
56	475	31.6	476	9	AI803690
57	473	31.4	481	9	AI673434
58	472.6	31.4	479	9	AW050614
59	467	31.0	474	12	BM712008
60	467	31.0	480	10	AW131630
61	456	30.3	457	9	AW050632
62	455	30.2	455	10	BE395404
63	455	30.2	456	9	AI378033
64	453.4	30.1	445	13	BU731451
65	452.8	30.1	645	12	BM825434
66	449	29.8	486	9	AA908673
67	448.8	29.8	567	14	W58472
68	446.4	29.7	946	10	BF581815
69	442.4	29.4	696	10	BE395162
70	438	29.1	444	10	AW452081
71	435.4	28.9	507	14	W58514
72	435	28.9	435	9	AI129923
73	434.8	28.9	438	9	AI214596
74	432.4	28.7	574	14	CF788313
75	430.6	28.6	723	12	BI837473
76	429	28.5	545	10	BF076147
77	428	28.4	430	9	AA448177

BI818880 603037354
BE796469 601589817
BF979853 602288368
BU624717 UI-H-FGI-
BG282361 602402863
AL568188 AL568188
CB851997 UI-CF-FNO
BM755987 K-EST0034
BM717452 UI-E-EJO-
BM848012 K-EST0127
BM674957 UI-E-EJO-
BF967768 602887657
AV654573 AV654573
BM723373 UI-E-EJO-
BE895089 601436007
BI464353 603204005
BG030249 602297847
BE617298 601441804
BM764342 K-EST0045
BI765897 603046146
BU539642 AGENCOURT
BM793929 K-EST0074
AL559036 AL559036
BM737591 K-EST0000
BE386240 601273447
BM742038 K-EST0014
BE869917 601446563
BM744024 K-EST0017
BM682228 UI-E-EJO-
BI761329 603044093
BG967186 602833867
BM675954 UI-E-EJO-
BG231973 naE34ql2.
BM743569 K-EST0016
CA424831 UI-H-FEI-
EX671560 BX671560
EG037061 602288368
BI546418 603188719
BM848396 K-EST0128
AW001572 wu34f05.x
BE302814 ba69c04.y
BI260323 602963342
BE962188 601655404
AL559035 AL559035
BF233600 602024015
W58459 zd25h06.s1
AI125225 qd87f01.x
BE931725 QV4-HT053
W58347 zd25b06.s1
AW974208 EST386311
BM708058 UI-E-CIL-
AI803690 tc19e02.x
AI673434 wf19a10.x
AW050614 wz19c10.x
BM712008 UI-E-EJO-
AW131630 xf32g02.x
AW050632 wz19e08.x
BE395404 601309913
AI378033 tes7g09.x
BU731451 UI-E-CIL-
BM825434 K-EST0097
AA908673 ol04d06.s
W58472 zd25b06.r1
BF581815 602099816
BE395162 601309505
AW452081 UI-H-BI3-
W58514 zd25h06.r1
AI129923 qc41d08.x
AI214596 tm28h07.x
CF788313 860251.MA
BI837473 603086791
BF076147 225644.MA
AA448177 zw83b12.s

```

FEATURES             Location/Qualifiers
     source            1..822
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="IMAGE:4780659"
                        /lab_host="DH10B (T1 phage-resistant)"
                        /clone_lib="NCI CGAP Skn3"
                        /note="Organ: skin; Vector: pCMV-Sport6; Site 1: NotI;
                        Site2: SalI; Cloned unidirectionally. Primer: Oligo df.
                        Average insert size 1.5kb. Library constructed by Life
                        Technologies. Note: this is a NCI_CGAP Library."

ORIGIN
Query Match          51.3%;   Score 772.4;   DB 12;   Length 822;
Best local similarity 99.5%;   Pred. No. 1.9e-126;
Matches 806;   Conservative 0;   Mismatches 1;   Indels 3;   Gaps 3;

```

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-x@mail.nih.gov
 Tissue Procurement: DCTD/DTP
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLC1269 row: j column: 06
 High quality sequence stop: 755.
 Location/Qualifiers
 1..829
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4560485"
 /tissue_type="renal cell adenocarcinoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 14"
 /note="Organ: kidney; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dr priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGACACAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 50.7%; Score 763.2; DB 12; Length 829;
 Best Local Similarity 97.2%; Pred. No. 7,8e-125;
 Matches 808; Conservative 0; Mismatches 19; Indels 4; Gaps 3;
 45 GCCGGCGCTCAGCTCTCGACCCCGCTGTGGGTAGTCCAGCGAGCGGCGGCGG 104
 Db 1 GCCGGCGCTCAGCTCTCGACCCCGCTGTGGGTAGTCCAGCGAGCGGCGGCGG 60
 105 TGGGCCCATGTCGAGCGCGCGGCTGAGCGGTGGCGGACCGGCTGGGCTGGTACGGG 164
 Db 61 TGGGCCCATGTCGAGCGCGCGGCTGAGCGGTGGCGGACCGGCTGGGCTGGTACGGG 120
 165 GGCCTCGGGGGGCTCGCGCGCGGCGGCTGGCGCGGCGGCTGGTCCAGCGAGGACTGAAGT 224
 Db 121 GGCCTCGGGGGGCTCGCGCGCGGCGGCTGGCGCGGCGGCTGGTCCAGCGAGGACTGAAGT 180
 225 GGTGGGCTGGCGCGGCTGTGGGCAACATCGAGAGCTGCTGCTGAATGAAGATGC 284
 Db 181 GGTGGGCTGGCGCGGCTGTGGGCAACATCGAGAGCTGCTGCTGAATGAAGATGC 240
 285 AGGCTACCCCGGACTTTGATCCCTACAGATGTGACCTATCAATGAAGAGCATCCT 344
 Db 241 AGGCTACCCCGGACTTTGATCCCTACAGATGTGACCTATCAATGAAGAGCATCCT 300
 345 CTCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTGTAGACATCTGCATCAACAATGC 404
 Db 301 CTCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTGTAGACATCTGCATCAACAATGC 360
 405 TGGCTTCGCGCGGCTCAGACCCCTGCTCTCAGGAGCAGCAGTGGTGGAGGACATGTT 464
 Db 361 TGGCTTCGCGCGGCTCAGACCCCTGCTCTCAGGAGCAGCAGTGGTGGAGGACATGTT 420
 465 CAATGTGAACGTGTGGCGGCTCAGCATCTGACACACGGGAGCCCTACAGTCCATGAAGGA 524
 Db 421 CAATGTGAACGTGTGGCGGCTCAGCATCTGACACACGGGAGCCCTACAGTCCATGAAGGA 480
 525 GCGGAATGTGAGCATGGGACATCAATTAACATCAATAGCATGTCTGGCCACCGAGTGT 584
 Db 481 GCGGAATGTGAGCATGGGACATCAATTAACATCAATAGCATGTCTGGCCACCGAGTGT 540
 585 ACCCTGTCTGTGACCCACTTATAGTGCCACCAAGTATGCCGTCACTGCGTGCACAGA 644
 Db 541 ACCCTGTCTGTGACCCACTTATAGTGCCACCAAGTATGCCGTCACTGCGTGCACAGA 600

645 GGGAGTGGGCAAGAGCTTCGGGAGGCCCAAGCCACATCCGAGCCAGCGTGCATCTCTCC 704
 Db 601 GGGAGTGGGCAAGAGCTTCGGGAGGCCCAAGCCACATCCGAGCCAGCGTGCATCTCTCC 660
 705 AGGTGTGGTGGAGA-CACAATTGCGCTTCAAACTCCACGACAGGACCCCTGAGAAGGCAG 763
 Db 661 AGGTGTGGTGGAGACCACAATTGCGCTTCAAACTCCACGACAGGACCCCTGAGAAGGCAT 720
 764 CTGCCACCTATGAGCAAAATGAAGTGTCTCAAAACCGAGGATGTGGCCGAGGCTGTATCT 823
 Db 721 CTG-CANCTATGAGCAAAATGAAGTGTCTCAAAACCGAGGATGTGGCCGAGGCTGTATCT 779
 824 ACGTCTCAGCACCCCGCCGACATCCAGATTGGAGACATCCAGATGAGGC 874
 Db 780 ACGTC--TCAGCACCCCGGACATTCGATTGGGGACTTCGATGAAGGC 828

RESULT 3
 BM471232
 LOCUS
 DEFINITION
 5', mRNA sequence.
 ACCESSION
 BM471232
 VERSION
 BM471232.1 GI:18520274
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 1010)
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC/DCTD/DTP
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLC1269 row: j column: 21
 High quality sequence stop: 561.
 Location/Qualifiers
 1..1010
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5563124"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_72"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 2 kb. Library constructed by Life
 Technologies."

FEATURES
 source
 1..1010
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5563124"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_72"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 2 kb. Library constructed by Life
 Technologies."

ORIGIN
 Query Match 48.7%; Score 732.4; DB 12; Length 1010;
 Best Local Similarity 92.2%; Pred. No. 2.1e-119;
 Matches 828; Conservative 0; Mismatches 61; Indels 9; Gaps 5;
 38 GAGAGCGCGCGGCTCAGCTCCTCCAGCCCGGATGAGCGGCTAGTCAGGAGGCGGACG 97
 Db 13 GAGAGCGCGCGGCTCAGCTCCTCCAGCCCGGATGAGCGGCTAGTCAGGAGGCGGACG 72
 98 GCGGGGCTGGGCGCCATGGCCAGCCCGGATGAGCGGCTGGCGGACCGGCTGGCGCTGG 157
 Db 73 GCGGGGCTGGGCGCCATGGCCAGCCCGGATGAGCGGCTGGCGGACCGGCTGGCGCTGG 132
 158 TGACGGGGGCTCGGGGGGATCGGCGCGCGGCTGGCGGCGGCTGGTCCAGGAGGAC 217

Db	133	TGACGGGGGCTCGGGGGGCAATCGGGGGCGCGTGGCCGGCCCTGGTCCAGCAGGGAC	192
Qy	218	TGAAGTGGTGGGTGGCGCGCACTGTGGGCAACATCGAGGAGCTGGCTGCTGAATGTA	277
Db	193	TGAAGTGGTGGGTGGCGCGCACTGTGGGCAACATCGAGGAGCTGGCTGCTGAATGTA	252
Qy	278	AGAGTGAGGGCTACCCCGGAGCTTTGATCCCTACAGATGTGACCTATCAATGAAGAGG	337
Db	253	AGAGTGAGGGCTACCCCGGAGCTTTGATCCCTACAGATGTGACCTATCAATGAAGAGG	312
Qy	338	ACATCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTGAGACATCTGCATCA	397
Db	313	ACATCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTGAGACATCTGCATCA	372
Qy	398	ACAATGTCTGGCTTGCGCCGGCTGCACACCTGCTCTCAGGCAGCACCGATGGTGGGAGG	457
Db	373	ACAATGTCTGGCTTGCGCCGGCTGCACACCTGCTCTCAGGCAGCACCGATGGTGGGAGG	432
Qy	458	ACATGTTCAATGTGAACGTGTGGCCCTCAGCATCTGCACACGGGAGCCTACCACTCCA	517
Db	433	ACATGTTCAATGTGAACGTGTGGCCCTCAGCATCTGCACACGGGAGCCTACCACTCCA	492
Qy	518	TGAAGGACGGNAATGTGACCGATGGGCAATCATTTAAACATCAATAGCATGCTGGCCAC	577
Db	493	TGAAGGACGGNAATGTGACCGATGGGCAATCATTTAAACATCAATAGCATGCTGGCCAC	552
Qy	578	GAGTGTACCCCTGTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTCACTGGCG	637
Db	553	GAGTGTACCCCTGTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTCACTGGCG	612
Qy	638	TGACAGAGGACTGAGGCAAGAGCTTGGGAGGCGCCAGACCCACATCCGAGCCAAGTGA	697
Db	613	TGACAGAGGACTGAGGCAAGAGCTTGGGAGGCGCCAGACCCACATCCGAGCCAAGTGA	672
Qy	698	TCTCTCCAGTGTGTGGAGACACAAATTCGCTTCAAACTCCACGACAGGACCCT-GAG	756
Db	673	TCTCTCCAGGGGGGTGGAGACCAATTCGGTTTCAAACTCCCGGACCAAGACCCCTGGAG	732
Qy	757	TAAGGACAGTGTG-CACCTATGAGCAATGAAGTGTCTCAAAACCCGAGGATGTGGCCGAG	813
Db	733	TAAGGAGCTGTGCCCCCTATGGACCAATGAAGTGTCTCCAAACCCGAGATGTGGCCGAG	792
Qy	814	GCTGTATTCTAGC-----TCCTCAGCACCCCGGCACACATCCAGATTGGAGACATCCAGAT	869
Db	793	GCTGTATTCTAGC-----TCCTCAGCACCCCGGCACACATCCAGATTGGAGAACCTCCCAAT	852
Qy	870	GA-GGCCCAACGAGCAGGTGAACCTAGTGAATGTGGGAGCTCTCTCTTCCCTCCCAACC	926
Db	853	GAGGGCCCAACGAGCAGGTGAACCTAGTGAATGTGGGAGCTCTCTCTTCCCTCCCAACC	910

RESULT 4	
BQ278156	
LOCUS	1071 bp mRNA linear EST 07-MAY-2002
DEFINITION	AGENCOURT_7061694 NIH_MGC_109 Homo sapiens cDNA clone IMAGE:S805128 5', mRNA sequence.
ACCESSION	BQ278156
VERSION	BQ278156.1 GI:20488364
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1. (bases 1 to 1071) NIH-MGC http://mgc.nci.nih.gov/ .
AUTHORS	National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE	Unpublished (1999)
JOURNAL	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov
COMMENT	Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

```

DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2043 row: n column: 09
High quality sequence stop: 603.
Location/Qualifiers
1..1071
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_image="5805128"
/tissue_type="teratocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 109"
/notes="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH MGC library."

```

ORIGIN	Query Match	47.8%;	Score 718.8;	DB 13;	Length 1071;
	Best Local Similarity	97.2%;	Pred. No. 5.4e-117;		
	Matches	774;	Conservative 0;	Mismatches 17;	Indels 5; Gaps 4;
Qy	17	GCAGTGGCGGGCGGCGCAGGAGCGGCGGGCGTCTCGACTCTCGACCCCGGTGTCGG	76		
Db	1	GCAGTGGCGGGCGGCGCAGGAGCGGCGGGCGTCTCGACTCTCGACCCCGGTGTCGG	60		
Qy	77	GCTAGTCCAGCGAGCGGGA CGGGCGGCGTGGGCCCATGGCCAGGCCCGGCATGAGCGGT	136		
Db	61	GCTAGTCCAGCGAGCGGGA CGGGCGGCGTGGGCCCATGGCCAGGCCCGGCATGAGCGGT	129		
Qy	137	GGCGGACCGGCTGGCGTGTGACGGGGGCTCTGGGGGGGCACTGGCGCGCGCGGTGGGCC	196		
Db	121	GGCGGACCGGCTGGCGTGTGACGGGGGCTCTGGGGGGGCACTGGCGCGCGCGGTGGGCC	180		
Qy	197	GGGCGCTGGTCCAGCAGGGACTGAAGGTGGTGGGGTGGCGCGGCACACTGTGGGCAACATCG	256		
Db	181	GGGCGCTGGTCCAGCAGGGACTGAAGGTGGTGGGGTGGCGCGGCACACTGTGGGCAACATCG	240		
Qy	257	AGGAGCTGGCTGCTGAATGTAAGAGTCAGAGGTACCCCGGGACTTTGATCCCTACAGAT	316		
Db	241	AGGAGCTGGCTGCTGAATGTAAGAGTCAGAGGTACCCCGGGACTTTGATCCCTACAGAT	300		
Qy	317	GTGACCTATCAAAATGAAGAGGACATCTCTCCATGTTCTCAGCTATCGCTTCTCAGCA	376		
Db	301	GTGACCTATCAAAATGAAGAGGACATCTCTCCATGTTCTCAGCTATCGCTTCTCAGCA	360		
Qy	377	GGGCTGTAGACATCTGCATCAACATGCTGGCTTGGCCGGCTCAGACACCTGCTCTCAG	436		
Db	361	GGGCTGTAGACATCTGCATCAACATGCTGGCTTGGCCGGCTCAGACACCTGCTCTCAG	420		
Qy	437	GCAGCACCAGTGGTTGGGAAGGACATGTTCAATGTGAACGTTGTCGGCCCTCAGCATCTGCA	496		
Db	421	GCAGCACCAGTGGTTGGGAAGGACATGTTCAATGTGAACGTTGTCGGCCCTCAGCATCTGCA	480		
Qy	497	CACGGGAAGCCTTACAGTCTCATGAAGAGCGGAATGTGGACCATGGGCATCATTAACA	556		
Db	481	CACGGGAAGCCTTACAGTCTCATGAAGAGCGGAATGTGGACCATGGGCATCATTAACA	540		
Qy	557	TCAATAGCATGTCCTGGCCACCGAGTGTACCCCTGCTGTGACCCACCTTCTATAGTGCCA	616		
Db	541	TCAATAGCATGTCCTGGCCACCGAGTGTACCCCTGCTGTGACCCACCTTCTATAGTGCCA	600		
Qy	617	CCAAATATGCCGTCACTGCGCTGACAGAGGACATGAGGCAAGCTTCGGGA-GGCCCCAG	675		
Db	601	CCAAATATGCCGTCACTGCGCTGACAGAGGACATGAGGCAAGCTTCGGAGGGCCAG	660		

```

676 ACCACATCCGAGCACCCTGATCTCTCCAGGTGT-GTGGAGACACAAATTCGCTTCAA 734
661 ACCACATCCGAGCACCCTGATCTCTCCAGGTGTGGTGGAGACAAATTCCTTCAA 720
735 ACTCACGA-CAAGACCCCTGAGAGGAGC--TGCCACCTATGAGCAAAATGAAGTGTCT 791
721 ACTCACGACCAAGGACCCCTGAGAGGAGCAGCTGGCCCCCTATGAACAAATGAAGTGTCC 780
792 CAACCCGAGGATGTG 807
781 TCMAACCCAGGAATG 796

RESULT 5
BI818880
LOCUS
DEFINITION
603037354F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5178397 5',
mRNA sequence.
BI818880
VERSION
BI818880.1 GI:15930430
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 940)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM11444 row: 1 column: 14
High quality sequence stop: 781.
FEATURES
Location/Qualifiers
1..940
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5178397"
/lab_host="DH10B"
/clone_lib="NIH_MGC_115"
/notes="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 Kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

ORIGIN
Query Match 46.6%; Score 701.4; DB 12; Length 940;
Best Local Similarity 93.6%; Pred. No. 6.5e-114;
Matches 864; Conservative 0; Mismatches 41; Indels 18; Gaps 12;

QY 1 CGCGGATCGACCAACGAGTC-GGCGGCGGCGGAGAGCGCGCGGCGTCAGCTC 59
DB 15 CGCGGATTCGACCAAGCAGGTCTGGCGGCGGCGGAGAGCGCGCGGCGTCAGCTC 74
QY 60 CTCGACCCCGTGTGCGGGCTAGTCCAGAGCGGAGCGGCGGCGGCGGCGGCGGCGG 119
DB 75 CTCGACCCCGTGTGCGGGCTAGTCCAGAGCGGAGCGGCGGCGGCGGCGGCGGCGG 134
QY 120 GCCCGGATGGAGCGGTGGCGGACCGGCTGGCGCTG-GTGACGGGCGGCGCTCGGGGGG-- 176

```

```

135 GCCCGGATGGAGCGGTGGCGGACCCGGCTGGCGTGTGTGACGGGGGCTCGGGGTGGC 194
177 -CATCGCGCGCGCGTGGCCCC-GGGCCCTGGTCCAGCAGGAGCTCAAGGTGTGGGCTGC 234
195 ATCTGGCTGGCGCGTGGCCCCCTGGTCCAGCAGGAGCTGAAGGTGTGGGCTGC 254
235 GCCCGACTGTGGGCAACATCGAGGAGCTGGCTGTGAATGTAAGAGTGCAGGCTACCCC 294
255 GCCCGACTGTGGGCAACATCGAGGAGCTGGCTGTGAATGTAAGAGTGCAGGCTACCCC 314
295 GGGACTTTGATCCCTCAGATGTGACCTATCAATGAGAGGACATCTCTCCATGTC 354
315 GGGACTTTGATCCCTCAGATGTGACCTATCAATGAGAGGACATCTCTCCATGTC 374
355 TCAGCTATCCGTTCTCAGCAGCGGGTGTAGACATCTGCATCAACAATGCTGGCTGGCC 414
375 TCAGCTATCCGTTCTCAGCAGCGGGTGTAGACATCTGCATCAACAATGCTGGCTGGCC 434
415 CGGCTGAC-ACCTGCTCTCAGCAGCACCAAGTGGTGGAGGACATGTTCAAT--GTG 471
435 CGGCTGACACACCTGCTCTCAGCAGCACCAAGTGGTGGAGGACATGTTCAATGTC 494
472 AACGTGTCGCCCTCAGCATCTGCACGCGGAGCCTACCACTCAACAATGCTGGCTGGCC 531
495 AGTTGCTGGCCCTCAGCATCTGCACGCGGAGCCTACCACTCAACAATGCTGGCTGGCC 554
532 GTGACGATGGGACATCAATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTG 591
555 GTGACGATGGGACATCAATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTG 614
592 TCTGTGACCACTTCTATAGTGCCACCAAGTATGCTGCTGCTGACAGAGGAGTGT 651
615 TCTGTGACCACTTCTATAGTGCCACCAAGTATGCTGCTGCTGACAGAGGAGTGT 674
652 AGGCAAGAGCTTCGGGAGGCCAGACCCACATCGAGCCAGCTGC-ATCTCTCCAGTGT 710
675 AGGCAAGAGCTTCGGGAGGCCAGACCCAGATTTCGAGCCAGTGTCAATCTCTCCAGTGT 734
711 GGTGGAGAC-ACAATTCGCTTCAAACTCCACAGACAGGACCCCTGAGAGGACAGTGC 769
735 GGGGAGACAAACAATTCGCTTCAAACTCCACAGACAGGACCCCTGAGAG--CAGTGGCA 792
770 CCTATGAGCAAAATGAA-GTGTCTCAAAACCGAGAGATGGCCGAGGTGTATCTACGTC 828
793 CCTATGAGCAAAATGAAAGGTGGCTCAAAACCGAGAGATGGCCGAGGTGTAAATCTACGTA 852
829 CTCAGCACCCCGCACACATCCAGATTGGAGATCCAGATGAGGCCACGAGCAGGTG 888
853 CTCAGNCCC---TGAACAATCAGAATGGAGATCCAAATTAAGGCCAGGGA-CAGGTG 908
889 ACCTAGTACTGTGGGAGTCTCT 911
909 ACATAGTAACTGTGGAACCTACT 931

RESULT 6
BE796469
LOCUS
DEFINITION
601589817F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3943948 5',
mRNA sequence.
BE796469
VERSION
BE796469.1 GI:10217667
KEYWORDS
EST.
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1063)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

```

Email: csaps-r@mail.nih.gov
 Tissue Procurement: DCTD/DTP
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: LiCM800 row: a column: 05
 High quality sequence stop: 782.

FEATURES
 Location/Qualifiers
 1..1063
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3943948"
 /tissue_type="small cell carcinoma"
 /cell_line="MGC3"
 /lab_hosts="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_7"
 /note="Organ: lung; Vector: pOTB7; Site: 1: XhoI; Site: 2: EcoRI; cDNA made by oligo-dr priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN
 Query Match 46.4%; Score 698.8; DB 10; Length 1063;
 Best Local Similarity 89.7%; Pred. No. 1.9e-113; Indels 11; Gaps 9;
 Matches 852; Conservative 0; Mismatches 87;

QY 1 CGCGGATCGGACCAAGCAGGTGCGCGCGCGGCGGAGAGCGCGCGGCGTCCAGCTCC 60
 Db 22 CGCGGATCGGACCAAGCAGGTGCGCGCGCGGCGGAGAGCGCGCGGCGTCCAGCTCC 81
 QY 61 TCGACCCCGGTGTCGGGTAGTCCN--CGGAGCGGAGCGGCGGCGTGGGCCATGGCCA 118
 Db 82 TCGACCCCGGTGTCGGGTAGTCCNATGCGATGCGGCGGCGGCGGCGTGGGCCATGGCCA 141
 QY 119 GCGCGGCGATCGAGCGGTGGCGC--GACCGGTGGCGGTGTCAGCGGCGCTCGGGGGCG 177
 Db 142 GCGCGGCGATCGAGCGGTGGCGCTGACCGGTGGCGGTGGCGGCGCTCGGGGGCG 201
 QY 178 ATCGCGCGCGCGTGGCGCGCGCTGTCGTCAGCAGGAGCTGAAGTGGTGGGTGCGCC 237
 Db 202 ATCGCGCGCGCGTGGCGCGCGCTGTCGTCAGCAGGAGCTGAAGTGGTGGGTGCGCC 261
 QY 238 CGCACTGGGCAACATCGAGAGTGGCTGCTGAATGTAAGTGCAGGCTACCGCGG 297
 Db 262 CGCACTGGGCAACATCGAGAGTGGCTGCTGAATGTAAGTGCAGGCTACCGCGG 321
 QY 298 ACTTTGATCCCTACAGATGTGACCTATCAATGAAGAGGACATCTCTCCATGTTCTCA 357
 Db 322 ACTTTGATCCCTACAGATGTGACCTATCAATGAAGAGGACATCTCTCCATGTTCTCA 381
 QY 358 GCTATTCGCTTTCAGCAGCGGTGTAGACATCTGCATCAACAATGCTGGGTGGCGCGG 417
 Db 382 GCTATTCGCTTTCAGCAGCGGTGTAGACATCTGCATCAACAATGCTGGGTGGCGCGG 441
 QY 418 CTTGACACCTGCTCTCAGGAGCAGCAGTGGTGGAGGACATCTCAATGTGAACGTG 477
 Db 442 CTTGACACCTGCTCTCAGGAGCAGCAGTGGTGGAGGACATCTCAATGTGAACGTG 501
 QY 478 CTGGGCGCTCAGCATCTGCACAGGAGCGCTACCACTGCATGAAGAGCGGAATGTGGAC 537
 Db 502 CTGGGCGCTCAGCATCTGCACAGGAGCGCTACCACTGCATGAAGAGCGGAATGTGGAC 561
 QY 538 GATGGGACATCATTAACATCAATAGCA--TGCTGGCGCACCGAGTGTACCCCTGCTGT 596
 Db 562 GATGGGACATCATTAACATCAATAGCATTTGCTGGCCACCGAGTGTACCCCTGCTGT 621

QY 597 GACCCACTTCTATAGTCCACCAAGTATGCGCTCACTGCGCTGACAGAGGAC--TGAGGC 655
 Db 622 GA--CCAATTCTATAGTGCCA--CAAGTATGCGCTCACTGCGCTGACAGAGGACTTGAGGC 679
 QY 656 AAGAGCTTCGGGAGGCGCCAGACCCACATCCGAGCCACGTGCA--TCTCTCCAGGTGTGGTG 714
 Db 680 AAGAGCTTCGGGAGGCGCCAGACCCACATCCGAGCCACGTGCA--TCTCTCCAGGTGTGGTG 739
 QY 715 GAGACACAATTCGCTTCAAACTCCACGACAGACCCCTGAGAAAGGAGCTGCCACCTAT 774
 Db 740 GAGACACAATTCGCTTCAAAAGTCCA--GACAAGAGACCCCTGAGAAAGGAGCTGGCAACTAT 798
 QY 775 GAGCAATGAAGTGTCTCAAAACCCGAGGATGTGCCGAGGCTGTATCTA--CGTCCTCA 832
 Db 799 AAGCAATGTAGTGTGTTAAACCCGAAATTTTGCCGAGGGATGGAATTAAGGTTCATCA 858
 QY 833 GCACCCCGCACACATCCAGATTCGAGACATCCAGATGAGGCCACCGAGGAGGTGACCT 892
 Db 859 GAAACCCCGCAAGATTCGGAATGTGGAAATTCACAATAATGCCAGGAGCGGACCT 918
 QY 893 AGTGACTGTGGAGCTCTCTCTCCCTCCACCCCTTCATGGCTTGCTC 942
 Db 919 ACGACCGTGGGGAAGCCCATATATCCACCAACTATGAGGGTGACCGC 968

RESULT 7
 BF979853/c
 LOCUS 602288368T1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4374184 3',
 DEFINITION mRNA sequence.
 ACCESSION BF979853
 VERSION BF979853.1 GI:12347068
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 679)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: csaps-r@mail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshinuki and Piero Carninci (RIKEN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LAM10037 row: 9 column: 17
 High quality sequence start: 24
 High quality sequence stop: 679.
 Location/Qualifiers
 1..679
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4374184"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_97"
 /note="Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site: 1: BamHI; Site 2: SalI-XhoI (gtcgag); Oligo-dr primed using primer 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

FEATURES
 Location/Qualifiers
 1..679
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4374184"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_97"
 /note="Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site: 1: BamHI; Site 2: SalI-XhoI (gtcgag); Oligo-dr primed using primer 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

ORIGIN


```

/clone="CS0DF035YM23"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

ORIGIN
Query Match 42.8%; Score 644.8; DB 9; Length 980;
Best Local Similarity 98.3%; Pred. No. 6.5e-104;
Matches 644; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 785 AGTGTCTCAAAACCCGAGGATGTGGCGAGCTGTATCTACGTCTCAGCACCACCCCGGCAC 844
Db 709 AGTGTCTCAAAACCCGAGGATGTGGCGAGCTGTATCTACGTCTCAGCACCACCCCGGCAC 650
QY 845 ACATCCAGATTGGAGACATCCAGATGAGGCCACGAGCAGCTGACCTAGTACTGTGG 904
Db 649 ACATCCAGATTGGAGACATCCAGATGAGGCCACGAGCAGCTGACCTAGTACTGTGG 590
QY 905 AGCTCTCTCTCCCTCCCAACCTTCATGGCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 964
Db 589 AGCTCTCTCTCCCTCCCAACCTTCATGGCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 530
QY 965 GATTTCTGGATCAGCGATACACCTCTCTGCTCCACACCCGAGCAGGCTGAGAAAATTT 1024
Db 529 GATTTCTGGATCAGCGATACACCTCTCTGCTCCACACCCGAGCAGGCTGAGAAAATTT 470
QY 1025 GTTTGAGATTTTATATCATCTGTCAAAATGCTTCAGTTGTAATGTGAAAATGGGCT 1084
Db 469 GTTTGAGATTTTATATCATCTGTCAAAATGCTTCAGTTGTAATGTGAAAATGGGCT 410
QY 1085 GGGGAAAGAGGTGTGCTTATGTTTACTGTTAACTGTTTGTGCTGCTGCTGCTGCTGCTG 1144
Db 409 GGGGAAAGAGGTGTGCTTATGTTTACTGTTAACTGTTTGTGCTGCTGCTGCTGCTGCTG 350
QY 1145 CACTTGGCTTGTGCTGCTCAGTGTCTCCCTTGATGAGGAAAGAGTTGGCCA 1204
Db 349 CACTTGGCTTGTGCTGCTCAGTGTCTCCCTTGATGAGGAAAGAGTTGGCCA 290
QY 1205 AAATCCCATCTCTTGGACCTCAACGCTCTGCTCAGGCTGGGTGGCAGAGGAGG 1264
Db 289 AAATCCCATCTCTTGGACCTCAACGCTCTGCTCAGGCTGGGTGGCAGAGGAGG 230
QY 1265 CCCTTCACTTATATCTGTTGTTATCCAGGGCTCCAGACTTCCTCTGCTGCTGCTGCTG 1324
Db 229 CCCTTCACTTATATCTGTTGTTATCCAGGGCTCCAGACTTCCTCTGCTGCTGCTGCTG 170
QY 1325 CTGCACCTCTCCCTTATCTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1384
Db 169 CTGCACCTCTCCCTTATCTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 110
QY 1385 GTCCCTCTCTGGGTCTATCCCTCCACTCTGACTCTGACTTGGCAGCAGAACACC 1439
Db 109 GTCCCTCTCTGGGTCTATCCCTCCACTCTGACTCTGACTTGGCAGCAGAACACTNC 55

RESULT 11
CB851997/c
LOCUS
DEFINITION CB851997 669 bp mRNA linear EST 22-APR-2003
UI-CF-FNO-aes-e-22-0-UI.s1 UI-CF-FNO Homo sapiens cDNA clone
CB851997
VERSION CB851997
KEYWORDS CB851997.1 GI:30046840
SOURCE EST.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 669)

```

AUTHORS TITLE

Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
79044477
8889548
Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
The following repetitive elements were found in this cDNA
sequence: 1-38, >POLY A#Simple_repeat (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES source

Location/Qualifiers
1..669
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-FNO-aes-e-22-0-UI"
/tissue_type="Human Lung Epithelial cells"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-FNO"
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-CF-FNO is a subtracted cDNA library derived from two
normalized Human lung epithelial cell libraries (ENI and
DUI) The library was subtracted according to according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. For additional information, contact:
bento-soares@uiowa.edu
TAG_SEQ=None found"

ORIGIN

Query Match 42.5%; Score 639.4; DB 14; Length 669;
Best Local Similarity 99.5%; Pred. No. 6e-103;
Matches 662; Conservative 0; Mismatches 1; Indels 2; Gaps 2;
QY 841 GCACATCCAGATTGGAGACATCCAGATGAGGCCACGAGCAGGTGACCTAGTACTG 900
Db 669 GCACATCCAGATTGGAGACATCCAGATGAGGCCACGAGCAGGTGACCTAGTACTG 610
QY 901 TGGAGCT 960
Db 609 TGGGA-CT 551
QY 961 TGTGATTTCTGGATCAGGGATACCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1020
Db 550 TGTGATTTCTGGATCAGGGATACCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 491
QY 1021 ATTGTTTGGATTTTATATCATCTGTCAAAATGCTTCAGTTGTAATGTGAAAATG 1080
Db 490 ATTGTTTGGATTTTATATCATCTGTCAAAATGCTTCAGTTGTAATGTGAAAATG 431
QY 1081 GGCTGGGAAAGAGGTGTGCTCCCTAAATGTTTAACTTGTAACTTGTCTTGTGCTGCTG 1140
Db 430 GGCTGGGAAAGAGGTGTGCTCCCTAAATGTTTAACTTGTAACTTGTCTTGTGCTGCTG 371
QY 1141 TGGGCACTTGGCTTGTCTGCTCTCAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1200
Db 370 TGGGCACTTGGCTTGTCTGCTCTCAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 312
QY 1201 GCCAAATCCCATCT 1260

Db 311 GCACAAATCCCACTCTTCTTGACCTCAAGTCTGTGGCTCAGGCTGGGTGCGAGG 252
 QY 1261 GAGCCCTTCACTTATATCTGTGTGTATATCCAGGGCTCCAGACTTCTCTCTGCTGC 1320
 Db 251 GAGCCCTTCACTTATATCTGTGTGTATATCCAGGGCTCCAGACTTCTCTCTGCTGC 192
 QY 1321 CCACCTGACCTCTCCCTTATCTATCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1380
 Db 191 CCACCTGACCTCTCCCTTATCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 132
 QY 1381 TCTGTGCT 1440
 Db 131 TCTGTGCT 72
 QY 1441 GGGCTGCT 1500
 Db 71 GGGCTGCT 12
 QY 1501 AAAAA 1505
 Db 11 AAAAA 7

RESULT 12
 BM755987
 LOCUS
 DEFINITION K-EST0034121 S1SNU5 Homo sapiens cDNA clone S1SNU5-38-A10 5', mRNA
 sequence.
 ACCESSION BM755987
 VERSION BM755987.1 GI:19085602
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 640)
 AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Kim,Y.S.
 TITLE 21C Frontier Korean EST Project 2001
 JOURNAL Unpublished (2002)
 COMMENT Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 38 row: A column: 10
 High quality sequence stop: 640.
 Location/Qualifiers
 1..640
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S1SNU5-38-A10"
 /sex="F"
 /tissue_type="Ascites"
 /cell_type="Lymphoblast-like"
 /lab_host="SNU-5"
 /lab_host="Top10F"
 /clone_lib="S1SNU5"
 /note="Organ: Stomach; Vector: pCNS; Site: 1: EcoRI;
 Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
 bacterial alkaline phosphatase (BAP) and then decapped
 with tobacco acid pyrophosphatase (TAP). The decapped
 intact mRNA was ligated with DNA-RNA linker including EcoR
 I site by treatment of T4 RNA ligase and the first strand
 cDNA was synthesized from oligo dt-selected mRNA by
 priming with dt-tailed vector. The dt-tailed vector was
 adjusted to have about 60nt. The cDNA vector was
 circularized with E. coli DNA ligase after digestion of
 EcoRI which site is also included in vector. An RNA strand

FEATURES

source
 1..640
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S1SNU5-38-A10"
 /sex="F"
 /tissue_type="Ascites"
 /cell_type="Lymphoblast-like"
 /lab_host="SNU-5"
 /lab_host="Top10F"
 /clone_lib="S1SNU5"
 /note="Organ: Stomach; Vector: pCNS; Site: 1: EcoRI;
 Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
 bacterial alkaline phosphatase (BAP) and then decapped
 with tobacco acid pyrophosphatase (TAP). The decapped
 intact mRNA was ligated with DNA-RNA linker including EcoR
 I site by treatment of T4 RNA ligase and the first strand
 cDNA was synthesized from oligo dt-selected mRNA by
 priming with dt-tailed vector. The dt-tailed vector was
 adjusted to have about 60nt. The cDNA vector was
 circularized with E. coli DNA ligase after digestion of
 EcoRI which site is also included in vector. An RNA strand

converted to a DNA strand by Okayama-Berg method. The
 obtained cDNA vectors were used for transposition of
 competent cells E. coli Top10F by electroporation method.
 The cDNA libraries constructed by this method are
 full-length enriched cDNA library.

ORIGIN

Query Match 42.4%; Score 638.4; DB 12; Length 640;
 Best Local Similarity 99.8%; Pred. No. 9.1e-103;
 Matches 639; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 350 TGTCTCAGCTATCCGTTCTCAGCACACCGGTGTAGACATCTGCATCAACATGCTGCT 409
 Db 1 TGTCTCAGCTATCCGTTCTCAGCACACCGGTGTAGACATCTGCATCAACATGCTGCT 60
 QY 410 TGGCCCGGCTGACACCTCTCTCAGGCAGCACCGTGTGAAGACATGTTCAATG 469
 Db 61 TGGCCCGGCTGACACCTCTCTCAGGCAGCACCGTGTGAAGACATGTTCAATG 120
 QY 470 TGAACGTCTGGCTCAGCATCTGCACACCGGAAGCTTACAGTCCATGAAGAGCGGA 529
 Db 121 TGAACGTCTGGCTCAGCATCTGCACACCGGAAGCTTACAGTCCATGAAGAGCGGA 180
 QY 530 ATGTGGACGATGGGCACATCAATTAACATAGCATCTTGGCCACCGAGTGTACCCC 589
 Db 181 ATGTGGACGATGGGCACATCAATTAACATAGCATCTTGGCCACCGAGTGTACCCC 240
 QY 590 TGTCTGTACCCACTTCTATAGTGCACCAAGTATGCCGTCACTCGCTGACAGAGGAC 649
 Db 241 TGTCTGTACCCACTTCTATAGTGCACCAAGTATGCCGTCACTCGCTGACAGAGGAC 300
 QY 650 TGAGGCAAGAGCTTCGGAGGCCACAGACCCACATCCAGAGCCACGTCATCTCTCCAGGTG 709
 Db 301 TGAGGCAAGAGCTTCGGAGGCCACAGACCCACATCCAGAGCCACGTCATCTCTCCAGGTG 360
 QY 710 TGGTGGAGACAAATTCGCTTCAAACTCCACGACAGGACCCCTGAGAAGGAGCTGCCA 769
 Db 361 TGGTGGAGACAAATTCGCTTCAAACTCCACGACAGGACCCCTGAGAAGGAGCTGCCA 420
 QY 770 CCTATGACCAATGAAGTGTCTCAAAACCGAGGATGTGCCGAGGCTGTATCTACGTCC 829
 Db 421 CCTATGACCAATGAAGTGTCTCAAAACCGAGGATGTGCCGAGGCTGTATCTACGTCC 480
 QY 830 TCAGCACCCCGCACACATCCAGATTGGAGACATCCAGATGAGGCCACGAGGAGTGA 889
 Db 481 TCAGCACCCCGCACACATCCAGATTGGAGACATCCAGATGAGGCCACGAGGAGTGA 540
 QY 890 CCTAGTACTGTGGAGCTCTCTCTTCCCTCCCAACCTTATGAGCTTGCCTTCCCTC 949
 Db 541 CCTAGTACTGTGGAGCTCTCTCTTCCCTCCCAACCTTATGAGCTTGCCTTCCCTC 600
 QY 950 TGGATTTAGGTGTGATTCTCTGGATCAGGGATACCACT 989
 Db 601 TGGATTTAGGTGTGATTCTCTGGATCAGGGATACCACT 640

RESULT 13

BM717452
 LOCUS
 DEFINITION UI-E-EJ0-ahn-c-08-0-UI.r1 UI-E-EJ0 Homo sapiens cDNA clone
 UI-E-EJ0-ahn-c-08-0-UI 5', mRNA sequence.
 ACCESSION BM717452
 VERSION BM717452.1 GI:19030710
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 649)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)

BM717452 649 bp mRNA linear EST 28-FEB-2002
 UI-E-EJ0-ahn-c-08-0-UI.r1 UI-E-EJ0 Homo sapiens cDNA clone
 UI-E-EJ0-ahn-c-08-0-UI 5', mRNA sequence.

```

MEDLINE
97044477
8889548
Contact: Soares, MB
COMMENT
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
The following repetitive elements were found in this cDNA
sequence: 603-644, >POLY_A$imple_repeat
Seq primer: M13 Reverse.
FEATURES
Location/Qualifiers
1..649
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-EJ0-ahn-c-08-0-UI"
/tissue_type="fetal eyes, lens, eye anterior segment,
optic nerve, retina, Retina Foveal and Macular, RPE and
Choroid"
/dev_stages="fetal and adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-EJ0"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-EJ0 is a subtracted cDNA library constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tags for this library are: fetal eyes,
AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,
AATCCGCGAT; optic nerve, CCATTAAAGTG; retina, CGCGG; Retina
Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This
library was created for the program, Gene Discovery in the
Visual System, supported by National Eye Institute (NEI)."
```

ORIGIN

```

Query Match 42.1%; Score 633; DB 12; Length 649;
Best Local Similarity 99.8%; Pred. No. 8.2e-102;
Matches 633; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 872 GGCCACGGAGCAGGTACCTAGTACTGTGGAGTCCCTTCCTCCCTCCACCTTCA 931
Db 1 GGCCACGGAGCAGGTACCTAGTACTGTGGAGTCCCTTCCTCCCTCCACCTTCA 60

QY 932 TGGCTTGCCTCCTCCCTCTGATTTTAGTGTGATTTCTGGATCAGGATACCACTTC 991
Db 61 TGGCTTGCCTCCTCCCTCTGATTTTAGTGTGATTTCTGGATCAGGATACCACTTC 120

QY 992 CTGTCCACCCCGACCGAGGGCTAGAAAAATTTGTTGAGATTTTATATCATCTTGTCA 1051
Db 121 CTGTCCACCCCGACCGAGGGCTAGAAAAATTTGTTGAGATTTTATATCATCTTGTCA 180

QY 1052 AATTGCTTCAGTGTAAATGTGAAAAATGGCTGGGAAAAGGAGTGGTCCCTAATTG 1111
Db 181 AATTGCTTCAGTGTAAATGTGAAAAATGGCTGGGAAAAGGAGTGGTCCCTAATTG 240

QY 1112 TTTTACTTGTAACTTCTTGTGGCCCTGGGACATGGCCCTTGTCTCTCTCAGTGT 1171
Db 241 TTTTACTTGTAACTTCTTGTGGCCCTGGGACATGGCCCTTGTCTCTCTCAGTGT 300
```

```

1172 CTTCCCTTTGACATGGGAAAGGAGTTGTGGCCAAAATCCCATCTTCTTGCACCTCAACG 1231
301 CTTCCCTTTGACATGGGAAAGGAGTTGTGGCCAAAATCCCATCTTCTTGCACCTCAACG 360
1232 TCTGTGGCTCAGGGCTGGGGTGGCAGAGAGGAGGCTTCCACCTTATATCTGTGTGTATC 1291
361 TCTGTGGCTCAGGGCTGGGGTGGCAGAGAGGAGGCTTCCACCTTATATCTGTGTGTATC 420
1292 CAGGGCTCCAGACTTCTCTCTGCTGCCCACTGCACCCCTCTCCCTTATCTATCTC 1351
421 CAGGGCTCCAGACTTCTCTCTGCTGCCCACTGCACCCCTCTCCCTTATCTATCTC 480
1352 CTTCTCGGCTCCCGAGCCAGCTCTTGGCTTCTTGTCCCTCTCTGGGGTCACTCCCTCCAT 1411
481 CTTCTCGGCTCCCGAGCCAGCTCTTGGCTTCTTGTCTCCTCTGGGGTCACTCCCTCCAT 540
1412 CTGACTCTGACTATGGCAGCAGAACACAGGSCCTGGCCCTGAGTGATTTTCATGTGTATCA 1471
541 CTGACTCTGACTATGGCAGCAGAACACAGGSCCTGGCCCTGAGTGATTTTCATGTGTATCA 600
1472 TTAAGAAAAAGAAATCGCAACCAAAAAA 1505
601 TTAAGAAAAAGAAATCGCAACCAAAAAA 634

RESULT 14
BM848012 627 bp mRNA linear EST 06-MAR-2002
LOCUS K-EST0127637 S13KMS5 Homo sapiens cDNA clone S13KMS5-43-A10 5',
DEFINITION mRNA sequence.
ACCESSION BM848012
VERSION BM848012.1 GI:19204411
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 627)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eosun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 43 row: A column: 10
High quality sequence stop: 627.
FEATURES
Location/Qualifiers
1..627
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S13KMS5-43-A10"
/tissue_type="myeloma"
/cell_line="KWS-5"
/lab_host="Top10F"
/clone_lib="S13KMS5"
/note="Vector: pCNS; Site 1: EcoRI; Site 2: NotI; The poly
(A)+ RNA was dephosphorylated with bacterial alkaline
phosphatase (BAP) and then decapped with tabacco acid
pyrophosphatase (TAP). The decapped intact mRNA was
ligated with DNA-RNA linker including EcoR I site by
treatment of T4 RNA ligase and the first strand cDNA was
synthesized from oligo dT-selected mRNA by priming with
dT-tailed vector. The dT-tailed vector was adjusted to
have about 60nt. The cDNA vector was circularized with E.
coli DNA ligase after digestion of EcoRI which site is
also included in vector. An RNA strand converted to a DNA
```


Hu,G., Gu,J., Chen,Z. and Han,Z.
Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)

21625106
11752456
COMMENT
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
source
Location/Qualifiers
1..651
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GLCDX005"
/tissue type="corresponding non cancerous liver tissue"
/dev stage="Adult"
/lab_host="SOIR"
/clone_lib="GLC"
/note="Vector: pBluescript sk(-); Site 1: EcoRI; Site_2:
XhoI"

ORIGIN
Query Match 41.2%; Score 620.8; DB 9; Length 651;
Best Local Similarity 97.4%; Pred. No. 1.2e-99; Indels 0; Gaps 0;
Matches 631; Conservative 0; Mismatches 17;

Qy 141 CGACCGCTGGCGCTGGTACGCGGGGCGCTCGGGGGGCATCGCGGGCGCGCGCGCGCGCGCG 200
Db 1 CGACCGCTGGCGCTGGTACGCGGGGCGCTCGGGGGGCATCGCGGGCGCGCGCGCGCGCGCG 60

Qy 201 CTGTGTCACGAGGAGTGAAGTGGTGGGCTGGCGCGCGCGCTGGCGCGCGCGCTGGCGCGCG 260
Db 61 CTGTGTCACGAGGAGTGAAGTGGTGGGCTGGCGCGCGCGCTGGCGCGCGCGCTGGCGCGCG 120

Qy 261 GTGTGCTGTGAATGATAGAGTGCAGGCTACCCCGGACTTTGATCCCTACAGATGTGA 320
Db 121 GTGTGCTGTGAATGATAGAGTGCAGGCTACCCCGGACTTTGATCCCTACAGATGTGA 180

Qy 321 CCTATCAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGG 380
Db 181 CCTATCAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGG 240

Qy 381 TGTAGACATCTGCATCAACAATGTGGCTTGGCGCGCGCGCTGACACCCCTGCTCTCAGGCG 440
Db 241 TGTAGACATCTGCATCAACAATGTGGCTTGGCGCGCGCGCTGACACCCCTGCTCTCAGGCG 300

Qy 441 CACCGTGTGGAGGACATGTTCAATGTGAACGTTGCGCCCTCAGCATCTGCACACG 500
Db 301 CACCGTGTGGAGGACATGTTCAATGTGAACGTTGCGCCCTCAGCATCTGCACACG 360

Qy 501 GGAAGCCTACCACTCCATGAAGAGCGGAAATGTGGACGATGGGCACATCATTAACATCAA 560
Db 361 GGAAGCCTACCACTCCATGAAGAGCGGAAATGTGGACGATGGGCACATCATTAACATCAA 420

Qy 561 TAGCATGTCTGGCCACCGAGTGTACCCCTGTCTGTGACCCACTTCTATAGTCCACCAA 620
Db 421 TAGCATGTCTGGCCACCGAGTGTACCCCTGTCTGTGACCCACTTCTATAGTCCACCAA 480

Qy 621 GTATGCCGTCACTGCGCTGACAGAGGAGCTGAGGCAAGAGCTTTCGGAGGCCCCAGACCCA 680
Db 481 GTATGCCGTCACTGCGCTGACAGAGGAGCTGAGGCAAGAGCTTTCGGAGGCCCCAGACCCA 540

Qy 681 CATCCGAGCCAGTGCATCTCTCAAGGTGTGGTGGAGACAAATTCGCCCTTCAAACTCCA 740
Db 541 CATCCGAGCCAGTGCATCTCTCAAGGTGTGGTGGAGACAAATTCGCCCTTAAACTTCA 600

Qy 741 CGACAAGGACCTGAGAGGCGAGCTGCCACTATGAGCAAAATGAAGTG 788
Db 601 CGACGAGGACCTGAGAGGCGAGCTGGCAACTATGAGCAAAAGGAAGTG 648

RESULT 18
BM723373
LOCUS
DEFINITION
UI-E-EJ0-aio-j-09-0-UI.r1 UI-E-EJ0 Homo sapiens cDNA clone
UI-E-EJ0-aio-j-09-0-UI 5', mRNA sequence.
BM723373
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

BM723373 658 bp mRNA linear EST 01-MAR-2002
UI-E-EJ0-aio-j-09-0-UI.r1 UI-E-EJ0 Homo sapiens cDNA clone
UI-E-EJ0-aio-j-09-0-UI 5', mRNA sequence.
BM723373
BM723373.1 GI:19044604
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 658)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
cDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Reverse.
Location/Qualifiers
1..658
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-EJ0-aio-j-09-0-UI"
/tissue type="fetal eyes, lens, eye anterior segment,
optic nerve, retina, Retina Foveal and Macular, RPE and
Choroid"
/dev stage="fetal and adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-EJ0"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-EJ0 is a subtracted cDNA library constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT7T3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tags for this library are: fetal eyes,
AGATACAGA; lens, CGATTAGGA; eye anterior segment,
AATGGCGCAT; optic nerve, CCATTAAAGT; retina, CCGCG;
Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This
library was created for the program, Gene Discovery in the
Visual System, supported by National Eye Institute (NEI)."

ORIGIN
Query Match 41.1%; Score 618.8; DB 12; Length 658;
Best Local Similarity 98.1%; Pred. No. 2.6e-99;
Matches 634; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

LOCUS B1464353 757 bp mRNA linear EST 21-AUG-2001
 DEFINITION 60320405F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5269728 5',
 mRNA sequence.
 ACCESSION B1464353
 VERSION B1464353.1 GI:15255009
 KEYWORDS EST.
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 757)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 DNA Sequencing by: Incyte Genomics, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM11680 row: j column: 01
 High quality sequence stop: 707.
 Location/Qualifiers
 1..757
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5269728"
 /lab_host="DH10B"
 /clone_lib="NIH MGC 97"
 /note="Organ: testis; Vector: pBluescriptR (modified
 pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
 (gtcgag); Oligo-dT primed using primer
 5'-TTTTTTTTTTTTTNN-3', size-selected for average
 insert size 2.2 kb and normalized to R0T 5. This is a
 primary library enriched for full-length clones and
 constructed using the Cap-trapper method (Carninci, in
 preparation). Library constructed by M. Brownstein
 (NIH/NHGRI, National Institutes of Health). Note: this is
 a NIH_MGC Library."
 ORIGIN
 Query Match 40.9%; Score 615.6; DB 12; Length 757;
 Best Local Similarity 98.9%; Pred. No. 9.6e-99;
 Matches 651; Conservative 0; Mismatches 4; Indels 3; Gaps 3;
 QY 1 CCGGGATCGGACCAAGCAGGTTCGGCGCGCGGCGGAGAGCGCGCGGCGTCACTCC 60
 DB 54 CCGGGATCGGACCAAGCAGGTTCGGCGCGCGGCGGAGAGCGCGCGGCGTCACTCC 113
 QY 61 TCGACCCCGTGTCTGGGCTAGTCCACGAGGCGGAGCGCGGCGGCGGCGGCGGCGGCGG 120
 DB 114 TCGACCCCGTGTCTGGGCTAGTCCACGAGGCGGAGCGCGGCGGCGGCGGCGGCGGCGG 173
 QY 121 CCGGGCATGAGCGGTGGCGCGGACCGGCTGGGCTGGTACGCGGCGGCGGCGGCGGCGGCGG 180
 DB 174 CCGGGCATGAGCGGTGGCGCGGACCGGCTGGGCTGGTACGCGGCGGCGGCGGCGGCGGCGG 233
 QY 181 GCGCGCGGCGTGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
 DB 234 GCGCGCGGCGTGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 293
 QY 241 ACTGTGGCAACATCGAGGAGCTGGCTGTGAATGTAAGAGTCCAGGCTACCCCGGAGCT 300
 DB 294 ACTGTGGCAACATCGAGGAGCTGGCTGTGAATGTAAGAGTCCAGGCTACCCCGGAGCT 353
 QY 301 TTGATCCCTTACAGATGTGACCTATCAATGAA- GAGGACATCTCTCCATGTTCTCAGC 359
 DB 354 TTGATCCCTTACAGATGTGACCTATCAATGAA- GAGGACATCTCTCTCCATGTTCTCAGC 413

QY 360 TATCCGTTCTCAGCACAGCGGTGTAGACATCTGCATCAACAATGCTGGCTTGGCCCGGCC 419
 DB 414 TATCCGTTCTCAGCACAGCGGTGTAGACATCTGCATCAACAATGCTGGCTTGGCCCGGCC 473
 QY 420 TGACACCCCTGCTCTCAGGCGAGCACCAGTGGTTGGAAGGACATGTTCAATGTGAACGTGCT 479
 DB 474 TGACACCCCTGCTCTCAGGCGAGCACCAGTGGTTGGAAGGACATGTTCAATGTGAACGTGCT 533
 QY 480 GGCCCTCAGCATCTGCACACGGGAAGCTTACCAGTCCATGAAGGAGCGGAATGTGGACGA 539
 DB 534 GGCCCTCAGCATCTGCACACGGGAAGCTTACCAGTCCATGAAGGAGCGGAATGTGGACGA 593
 QY 540 TGGGCACA-TCATTAAACATCAATAGCATGTCTGGCCACCGAGTGTACCC- TGTCTGTG 597
 DB 594 TGGGCACATTCATTAAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTTGTCTGTG 653
 QY 598 ACCCACTTCTATAGTGGCCACCAAGTATGCCGTCTACTGCGTGACAGAGGAGTGAAGC 655
 DB 654 AACCAATTCTATAGTGGCCACCAAGTATGCCGTCTACTGCGTGACAGAGGAGTGAAGC 711
 RESULT 21
 LOCUS B030249 932 bp mRNA linear EST 24-JAN-2001
 DEFINITION 602297547F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:4392011 5',
 mRNA sequence.
 ACCESSION B030249
 VERSION B030249.1 GI:12419347
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 932)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: DCTD/DTF
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM10083 row: n column: 12
 High quality sequence stop: 601.
 Location/Qualifiers
 1..932
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4392011"
 /tissue_type="mammary adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 87"
 /note="Organ: Breast; Vector: PCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.383 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."
 ORIGIN
 Query Match 40.9%; Score 615.6; DB 10; Length 932;
 Best Local Similarity 94.7%; Pred. No. 9.5e-99;
 Matches 658; Conservative 0; Mismatches 35; Indels 2; Gaps 2;
 QY 264 GGTGCTGTAATGTAAGTGCAGGCTACCCCGGAGCTTTGATCCCTTACAGATGTGACCT 323
 DB 10 GCTGCTGTAATGTAAGTGCAGGCTACCCCGGAGCTTTGATCCCTTACAGATGTGACCT 69

```

324 ATCAATGAAGAGGACATCTCTCCATGTTCTCAGCTATCCGTTCTCAGCAGCGGTGT 383
    |||
    |||
70 ATCAATGAAGAGGACATCTCTCCATGTTCTCAGCTATCCGTTCTCAGCAGCGGTGT 129
    |||
    |||
384 AGACATCTGCATCAACATGCTGGCTTGGCCCGGCTGACACCCCTGCTCTCAGGAGCAG 443
    |||
    |||
130 AGACATCTGCATCAACATGCTGGCTTGGCCCGGCTGACACCCCTGCTCTCAGGAGCAG 189
    |||
    |||
444 CAGTGGTGGAGGAGCATGTTCAATGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 503
    |||
    |||
190 CAGTGGTGGAGGAGCATGTTCAATGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 249
    |||
    |||
504 AGCTTACCAGTCCATGAAGGAGCGGAATGTGGACGATGGGACATCAATTAACATCAATAG 563
    |||
    |||
250 AGCTTACCAGTCCATGAAGGAGCGGAATGTGGACGATGGGACATCAATTAACATCAATAG 309
    |||
    |||
564 CATGCTTGGCCACCGAGTGTACCCCTGTGTGACCCCACTTCTATAGTGGCCACCAAGTA 623
    |||
    |||
310 CATGCTTGGCCACCGAGTGTACCCCTGTGTGACCCCACTTCTATAGTGGCCACCAAGTA 369
    |||
    |||
624 TGCCGTCACTGGCTGACAGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 683
    |||
    |||
370 TGCCGTCACTGGCTGACAGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 429
    |||
    |||
684 CCGAGCCACGTCATCTCTCCAGTGTGGTGGAGACAAATTCGCTTCAAACCTCCACGA 743
    |||
    |||
430 CCGAGCCACGTCATCTCTCCAGTGTGGTGGAGACAAATTCGCTTCAAACCTCCACGA 489
    |||
    |||
744 CAAGGACCTGAGAGGAGCTGCCCTATGAGCAATGAAGTGTCTCAAACCGGAGGA 803
    |||
    |||
490 CAAGGACCTGAGAGGAGCTGCCCTATGAGCAATGAAGTGTCTCAAACCGGAGGA 549
    |||
    |||
804 TGTGGCCGAGGCTGTTATCTACGTCCTCAGACACCCCGGACACATTCAGATTTGAGACAT 863
    |||
    |||
550 TGTGG-CGAGGCTGTTATCTACGTCCTCAGACACCCCGGACACATTCAGATTTGAGACAT 608
    |||
    |||
864 CCAGATGAGGCCACCGAGGAGTGAAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAG 923
    |||
    |||
609 CAGATGAGGCCACCGAGGAGTGAAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAG 667
    |||
    |||
924 ACCCTTATGCTTGCCTTCCCTCTGGATTTTA 958
    |||
    |||
668 ACCTTTAGTGTGCTCCGCTCTGGATTTAGGTGTGA 702
    |||
    |||

```

```

RESULT 22
BE617298
LOCUS
DEFINITION
60141804F1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3846168 5',
mRNA sequence.
BE617298
BE617298.1 GI:9888236
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM9558 row: 9 column: 01
High quality sequence stop: 698.
Location/Qualifiers

```

```

source
1..923
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3846168"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_65"
/note="Organ: colon; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.8 kb. Library constructed by Life
Technologies."
ORIGIN
Query Match 40.1%; Score 603.4; DB 10; Length 923;
Best Local Similarity 99.8%; Pred. No. 1.4e-96;
Matches 604; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 332 AAGAGGACATCTCTCCATGTTCTCAGCTATCCGTTCTCAGCAGCGGTGTAGACATCT 391
Db 1 AAGAGGACATCTCTCCATGTTCTCAGCTATCCGTTCTCAGCAGCGGTGTAGACATCT 60
QY 392 GCATCAACAATGCTGGCTTGGCCCGGCTGACACCCCTGCTCTCAGGAGCAGCGGTGTAG 451
Db 61 GCATCAACAATGCTGGCTTGGCCCGGCTGACACCCCTGCTCTCAGGAGCAGCGGTGTAG 120
QY 452 GGAAGGACATGTTCAATGTGAAGTGTGGCCCTCAGCATCTGCACACGGGAGCCTACC 511
Db 121 GGAAGGACATGTTCAATGTGAAGTGTGGCCCTCAGCATCTGCACACGGGAGCCTACC 180
QY 512 AGTCCATGAGGAGCGGAATGTGACCATGGGACATCAATTAACATCAATAGCATGTCTG 571
Db 181 AGTCCATGAGGAGCGGAATGTGACCATGGGACATCAATTAACATCAATAGCATGTCTG 240
QY 572 GCCACCGAGTGTACCCCTGCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCCTCA 631
Db 241 GCCACCGAGTGTACCCCTGCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCCTCA 300
QY 632 CTGCGCTGACAGAGGAGTGAAGGAGCTTGGGAGGCGCCAGACCCACATCCGAGCCA 691
Db 301 CTGCGCTGACAGAGGAGTGAAGGAGCTTGGGAGGCGCCAGACCCACATCCGAGCCA 360
QY 692 CGTGCATCTCTCCAGGTTGGTGGAGACAAATTCGCTTCAAACCTCCAGCAAGGAGC 751
Db 361 CGTGCATCTCTCCAGGTTGGTGGAGACAAATTCGCTTCAAACCTCCAGCAAGGAGC 420
QY 752 CTGAGAGGAGGAGTGCACCCCTATGAGCAATGAAGTGTCTCAAACCGAGGATGTGGCCG 811
Db 421 CTGAGAGGAGGAGTGCACCCCTATGAGCAATGAAGTGTCTCAAACCGAGGATGTGGCCG 480
QY 812 AGGCTGTTATCTAGTCTCAGCAGCCCGGACACATCCAGATTTGGAGACATCCAGATGA 871
Db 481 AGGCTGTTATCTAGTCTCAGCAGCCCGGACACATCCAGATTTGGAGACATCCAGATGA 540
QY 872 GGCCTCAGGAGGAGTGAAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGA 931
Db 541 GGCCTCAGGAGGAGTGAAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGA 600
QY 932 TGGCT 936
Db 601 TGGTT 605

```

```

RESULT 23
BM764342
LOCUS
DEFINITION
K-EST0045842 SISNU5 Homo sapiens cDNA clone SISNU5-39-A06 5', mRNA
sequence.
ACCESSION
BM764342
VERSION
BM764342.1 GI:19093957
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 602)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
Unpublished (2002)
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 39 row: A column: 06
High quality sequence stop: 602.

[illegible]

```

1. 602
Location: /usr/local/lib
/organelle="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cloned="SISNU5-39-A06"
/sex="F"
/tissue_type="Ascites"
/cell_type="Myeloblast-like"
/cell_line="G8U-5"
/lab_host="Top10f"
/cloned_lib="SISNU5"
/note="Organ: Stomach; Vector: Site 2; Noti; The poly (A) + R bacterial alkaline phosphatase with tabacco acid pyrophosphatase with intact mRNA was ligated with intact mRNA was ligated with T4 RNA site by treatment of T4 RNA cDNA was synthesized from oligo priming with dr-tailed vector constructed to have about 60nt. circularized with E. coli DNA EcoRI which site is also included converted to a DNA strand by obtained cDNA vectors were used competent cells E. coli Top10f. The cDNA libraries constructed from full-length enriched cDNA libraries."

```

ORIGIN

Query Match	40.0%;	Score 602;	DB 12;	Length 602;
Best Local Similarity	100.0%;	Prod. No. 2.5e-96;		
Matches 602;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	350	TGTTCTCAGCTATCGTTTCTCAGACACGGGTGATAGCATCTGCATCAACAATGCTGGCT	409	
Db	1	TGTTCTCAGCTATCGTTTCTCAGACACGGGTGATAGCATCTGCATCAACAATGCTGGCT	60	
Qy	410	TGGCCCGGCTTGACACCTGTCTCTCAGGAGCACCACTGGTTTGGGAAGGACATGTTTCAATG	469	
Db	61	TGGCCCGGCTTGACACCTGTCTCTCAGGAGCACCACTGGTTTGGGAAGGACATGTTTCAATG	120	
Qy	470	TGMACTGTCTGGCCCTCAGCATCTGCACACGGGAAGCCTACCAAGTCCATGAAGGAGCGGA	529	
Db	121	TGMACTGTCTGGCCCTCAGCATCTGCACACGGGAAGCCTACCAAGTCCATGAAGGAGCGGA	180	
Qy	530	ATGTGGACGATGGGCACATCAATTAACATCAATAGCATGTCTGGCCACCGAGTGTTTACCCC	589	
Db	181	ATGTGGACGATGGGCACATCAATTAACATCAATAGCATGTCTGGCCACCGAGTGTTTACCCC	240	
Qy	590	TGTCGTGACCCACTTCTATAGTGGCCACCAAGTATGCCGTCACTGGCTGCAGAGGAGAC	649	
Db	241	TGTCGTGACCCACTTCTATAGTGGCCACCAAGTATGCCGTCACTGGCTGCAGAGGAGAC	300	
Qy	650	TGAGGCAAGAGCTTCGGGAGGGCCAGACCCACATCCGAGGCCACGTGCATCTCTCCAGGTG	709	

Db	301	TGAGGCAAGAGCTTCGGAGGCCCGACACCATCCGAGCCACGTCATCTCTCCAGGTG	360
Qy	710	TGTTGGAGACACAATTCGCTTTAAACTCCACGACAAGGACCTTGAGAAAGCAGCTGCCA	769
Db	361	TGTTGGAGACACAATTCGCTTTAAACTCCACGACAAGGACCTTGAGAAAGCAGCTGCCA	420
Qy	770	CCATATGAGCAAAATGAAAGTGCTCAAAACCCGAGGATGTGGCCGAGGCTCTTATCTACGTCC	829
Db	421	CCATATGAGCAAAATGAAAGTGCTCAAAACCCGAGGATGTGGCCGAGGCTCTTATCTACGTCC	480
Qy	830	TCAGCACCCCGCACACATCCAGATTGGAGACATCCAGATGAGGCCCAACGAGGAGGTGA	889
Db	481	TCAGCACCCCGCACACATCCAGATTGGAGACATCCAGATGAGGCCCAACGAGGAGGTGA	540
Qy	890	CCATAGTCACTGTGGGAGCTCTCTCCCTCCACACCTTCATGGCTTGCCTTCCTGCCTC	949
Db	541	CCATAGTCACTGTGGGAGCTCTCTCTCCCTCCACACCTTCATGGCTTGCCTTCCTGCCTC	600
Qy	950	TG 951	
Db	601	TG 602	

RESULT 24

RECEIVED
BI765897
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCES
AUTHORS
TITLE
JOURNAL
COMMENT

TABLE 11.10.3

FEATURES

source

source

ORIGIN

Query Match	39.9%;	Score 600.8;	DB 12;	Length 804;
Best Local Similarity	94.2%;	Pred. No. 4e-96;		

```
Matches 680; Conservative 0; Mismatches 32; Indels 10; Gaps 5;
QY 1 CGCGATCGACCAAGCAGGTGCGCGCGCGGAGAGAGCGCGCGCGCTCAGCTCC 60
Db 83 CGCGATCGACCAAGCAGGTGCGCGCGCGGAGAGAGCGCGCGCGCTCAGCTCC 142
QY 61 TCGACCCCGTGTGCGGCTAGTCCAGCGAGCGGAGCGCGCGCGCTGCGCGCGCT 120
Db 143 TCGACTCCCGTGTGCGGCTAGTCCAGCGAGCGGAGCGCGCGCGCTGCGCGCGCT 202
QY 121 CCCGCGATGAGCGGTGCGCGCGGCTGCGCGCGCTGCGCGCGCTGCGCGCGCT 180
Db 203 CCCGCGATGAGCGGTGCGCGCGGCTGCGCGCGCTGCGCGCGCTGCGCGCGCT 262
QY 181 GCGCGCGCGCTGCGCGCGGCTGCGCGCGGCTGCGCGCGCTGCGCGCGCT 300
Db 263 GCGCGCGCGCTGCGCGCGGCTGCGCGCGGCTGCGCGCGCTGCGCGCGCT 360
QY 241 ACTGTGGGCAACATCGAGGAGCTGCTGCTGAATGTGAAGTGCAGGCTACCCCGGACT 300
Db 323 ACTGTGGGCAACATCGAGGAGCTGCTGCTGAATGTGAAGTGCAGGCTACCCCGGACT 382
QY 301 TTGATCCCGCTCAGATGTGACCTATCAATGAAGAGGACATCTCTCCATGTTCTCAGCT 360
Db 383 TTGATCCCGCTCAGATGTGACCTATCAATGAAGAGGACATCTCTCCATGTTCTCAGCT 442
QY 361 ATCCGTTCTCAGCAGCGGTGAGACATCTGCATCAACAAATGCTGGCTTGGCCCGGCT 420
Db 443 ATCCGTTCTCAGCAGCGGTGAGACATCTGCATCAACAAATGCTGGCTTGGCCCGGCT 502
QY 421 GACACCTGCTCTCAGGAGCAGCAGTGTGTTGGAAGGACATGTTCAATGTGAACGTGCTG 480
Db 503 GACACCTGCTCTCAGGAGCAGCAGTGTGTTGGAAGGACATGTTCAATGTGAACGTGCTG 562
QY 481 GCCCTCAGCATCTGACACGCGGAGCTTACAGTCCATGAAGAGCGGAATGTGG--ACG 538
Db 563 GCCCTCAGCATCTGACACGCGGAGCTTACAGTCCATGAAGAGCGGAATGTGGGAACG 622
QY 539 ATGGGCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 597
Db 623 ATGGGCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 682
QY 598 ACCCATCTCTATA--GTGCCACCAAGTATGCGCTCACTGCGC--TGACAGAGGAGCTGA 652
Db 683 ACCCATCTCTATAAGTGCACCAAGTATGCGCTCACTGCGC--TGACAGAGGAGCTGA 742
QY 653 GGCAGAG--CTTCGGAGGCGGAGACCCACATCGAGCCACGTGCTCTCCAGGTGT 710
Db 743 GGCAGAGACTTCGGGAAGGCCAAACCAACATCGAGACAAAGTGCATCTCTCCAGGTGT 802
QY 711 GG 712
Db 803 TG 804

RESULT 25
BUS39642
LOCUS
DEFINITION
AGENCOURT 10224058 NIH_MGC_107 Homo sapiens cDNA clone
IMAGE:6570429 5', mRNA sequence.
ACCESSION
BUS39642
VERSION
BUS39642.1 GI:22850083
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 937)
AUTHORS
NIH-MGC http://mgi.nci.nih.gov/
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
```

Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LLCM2759 row: e column: 21
High quality sequence stop: 527.
Location/Qualifiers
1. .937
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6570429"
/tissue_type="adenocarcinoma, cell_line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_107"
/note="Organ: breast; Vector: pOTB7; Site 1: EcoRI; Site 2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

```
Query Match      39.6%; Score 596.6; DB 13; Length 937;
Best Local Similarity 95.4%; Pred. No. 2.2e-95;
Matches 625; Conservative 0; Mismatches 29; Indels 1; Gaps 1;
QY 135 GTGCGCGACCGCTGGCGCTGGTGAACGGGGGCGCTCGGGGGGCGCATCGCGGGCGCTGCG 194
Db 1 GTGCGCGACCGCTGGCGCTGGTGAACGGGGGCGCTCGGGGGGCGCATCGCGGGCGCTGCG 60
QY 195 CCGGGCCCTGTCCAGCAGGAGCTGAAGTGGTGGGTGCGCCCGCACTGTGGGCAACAT 254
Db 61 CCGGGCCCTGTTCAGCAGGAGCTGAAGTGGTGGGTGCGCCCGCACTGTGGGCAACAT 120
QY 255 CGAGGAGCTGCTGCTGAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 314
Db 121 CGAGGAGCTGCTGCTGAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 180
QY 315 ATGTGACCTATCAATGAAGAGGACATCTCTCCATGTTCTCAGCTATCCGTTCTCAGCA 374
Db 181 ATGTGACCTATCAATGAAGAGGACATCTCTCCATGTTCTCAGCTATCCGTTCTCAGCA 240
QY 375 CAGCGGTGTAGACATCTGCATCAACAAATGCTGGCTTGGCCCGGCTGACACCCCTGCTCTC 434
Db 241 CAGCGGTGTAGACATCTGCATCAACAAATGCTGGCTTGGCCCGGCTGACACCCCTGCTCTC 300
QY 435 AGGCGACACAGTGTGTTGGAAGGACATGTTCAATGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 494
Db 301 AGGCGACACAGTGTGTTGGAAGGACATGTTCAATGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 360
QY 495 CACACGGGAAGCTTACCAAGTCCATGAAGAGCGGAATGTGAGAGTGGGACATCATTA 554
Db 361 CACACGGGAAGCTTACCAAGTCCATGAAGAGCGGAATGTGAGAGTGGGACATCATTA 420
QY 555 CATCAATAGATGTCTGGCCACCGAGTGTACCCCTGTCTGTGACCCCACTTCTATAGTGC 614
Db 421 CATCAATAGATGTCTGGCCACCGAGTGTACCCCTGTCTGTGACCCCACTTCTATAGTGC 480
QY 615 CACCAAGTATGCGCTCACTGCGCTGACAGAGGAGCTGAGGCAAGAGCTTCGGAGAGCCCA 674
Db 481 CACCAAGTATGCGCTCACTGCGCTGACAGAGGAGCTGAGGCAAGAGCTTCGGAGAGCCCA 540
QY 675 GACCCACATCCGAGCCACGCTGCTCTCAGGTGTGGTGGAGACAAATT--CGCCTTCA 733
Db 541 GACCCACATCCGAGCCACGCTGCTCTCAGGTGTGGTGGAGACAAATT--CGCCTTCC 600
QY 734 AACTCCACAGCAAGGACCTTGAGAGGAGCTGCCACCTATGAGCAAAATGAAGT 788
```


Db	601	AACTCCAGAACAGGACCCCTGTAGAAAGGCTTCTGCCACCTTAGAACCAAATG	655
RESULT 26			
BM793929			
LOCUS		691 bp mRNA linear	EST 05-MAR-2002
DEFINITION		K-EST0074920 S22SNU16n1 Homo sapiens cDNA clone S22SNU16n1-32-C08	
5', mRNA sequence.			
BM793929			
BM793929.1		GI:19142161	
EST.			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
1		(bases 1 to 691)	
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,			
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and			
Kim,Y.S.			
21C Frontier Korean EST Project 2001			
Unpublished (2002)			
Contact: Kim YS			
Genome Research Center			
Korea Research Institute of Bioscience & Biotechnology			
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea			
Tel: +82-42-860-4470			
Fax: +82-42-860-4409			
Email: yongsung@mail.kribb.re.kr			
Plate: 32 row: C column: 08			
High quality sequence stop: 691.			
Location/Qualifiers			
1..691			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="S22SNU16n1-32-C08"			
/sex="F"			
/tissue_type="Ascites"			
/cell_type="Lymphoblast-like"			
/cell_line="SNU-16"			
/lab_host="DH108"			
/clone_lib="S22SNU16n1"			
/note="Organ: Stomach; Vector: pT73-Pac; Site:1: EcoRI;			
Site:2: NotI; The S22SNU16 library was contributed by the			
Soares laboratory and it was constructed as described by			
Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome			
Research 6(9): 791-806. RNA was prepared from harvested			
cells of SNU-16 culture. SNU-16 cell was obtained from			
Korean Cell Line Bank (KCLB). SNU-16 was established from			
ascitic fluids of Korean patients by Park J.G. et al.			
(1990), Cancer Res 50: 2773-2780."			
ORIGIN			
Query Match			
Best Local Similarity			
Matches			
690; Conservative			
0; Mismatches			
1; Indels			
93; Gaps			
1;			
QY	472	AAGCTGCTGCGCCCTCAGCATCTGCACACGGGAAGCCTACGAGTCCATGAAGGACCGGAAT	531
Db	1	AAGCTGCTGCGCCCTCAGCATCTGCACACGGGAAGCCTACGAGTCCATGAAGGACCGGAAT	60
QY	532	GTGGACGATGGGCACATCATTAACATCAATAGATGTCTGGCCACCGAGTGTACCCCTG	591
Db	61	GTGGACGATGGGCACATCATTAACATCAATAGATGTCTGGCCACCGAGTGTACCCCTG	120
QY	592	TCTGTGACCCACTTCTATAGTGGCCACCAAGTATGCGTCACTCGCTGCAGAGGAGCTG	651
Db	121	TCTGTGACCCACTTCTATAGTGGCCACCAAGTATGCGTCACTCGCTGCAGAGGAGCTG	180
QY	652	AGGCAAGAGCTTCGGGAGGGCCAGACCCACATCCGAGCCACGTGTCCTCCAGGTGTG	711
Db	181	AGGCAAGAGCTTCGGGAGGGCCAGACCCACATCCGAGCCAC	----- 221
QY	712	GTGGAGACACAATTCGCTTCAAACTCCACGACAAAGGACCCCTGAGAGGACGAGTGCAC	771
Db	222	-----	221
QY	772	TATGAGCAAAATGAAGTGTCTCAAAACCGAGAGATGTGCCGAGGCTGTATCTACGTCTC	831
Db	222	-----GTGCTCTCAAAACCGAGAGATGTGCCGAGGCTGTATCTACGTCTC	267
QY	832	AGCAACCCCGCACATCCAGATTGGAGACATCCAGATGAGGCCACGAGAGAGGTGACC	891
Db	268	AGCAACCCCGCACATCCAGATTGGAGACATCCAGATGAGGCCACGAGAGAGGTGACC	327
QY	892	TAGTGACTGTGGGAGCTCTCTCTCCCTCCCAACCTTTCATGGCTTGCTCTCTGCTCTG	951
Db	328	TAGTGACTGTGGGAGCTCTCTCTCCCTCCCAACCTTTCATGGCTTGCTCTCTGCTCTG	387
QY	952	GATTTTAGTGTGTTGATTTCTGGATCAGGGGATACCACTCTCTGTCACACCCGACACAGG	1011
Db	388	GATTTTAGTGTGTTGATTTCTGGATCAGGGGATACCACTCTCTGTCACACCCGACACAGG	447
QY	1012	GGCTAGAAAAATTTGTTGAGATTTTATATCATCTGTCAAAATGCTTCAGTTGTAAATG	1071
Db	448	GGCTAGAAAAATTTGTTGAGATTTTATATCATCTGTCAAAATGCTTCAGTTGTAAATG	507
QY	1072	TGAAAAATGGCGCTGGGGAAGAGGAGTGTGTCCTTAATTTTACTTTGTTAACTTGTTC	1131
Db	508	TGAAAAATGGCGCTGGGGAAGAGGAGTGTGTCCTTAATTTTACTTTGTTAACTTGTTC	567
QY	1132	TTGTGCCCTGGGCACATTTGGCTTTGTCTGTCTCAGTGTCTTCCCTTTGACATGGGAAA	1191
Db	568	TTGTGCCCTGGGCACATTTGGCTTTGTCTGTCTCAGTGTCTTCCCTTTGACATGGGAAA	627
QY	1192	GGAGTTGTGGCCAAAATCCCATCTTCTTGACACCTCAACGCTGTGGCTCAGGGCTGGG	1251
Db	628	GGAGTTGTGGCCAAAATCCCATCTTCTTGACACCTCAACGCTGTGGCTCAGGGCTGGG	687
QY	1252	TGGC	1255
Db	688	TGGC	691
RESULT 27			
AL559036			
LOCUS		1022 bp mRNA linear	EST 31-MAY-2003
DEFINITION		Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED	
AL559036			
AL559036.2		GI:31283169	
EST.			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
On Feb 15, 2001 this sequence version replaced gi:12904138.			
Contact: Genoscope			
Genoscope - Centre National de Sequencage			
Bp 191 91006 EVRY cedex - France			
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr			
was normalized. Library was constructed by Life Technologies, a			
division of Invitrogen. This sequence belongs to sequence cluster			
8845.f For more information about this cluster, see			
http://www.genoscope.cns.fr/			
cgi-bin/cluster.cgi?seq=CS0DJ010CG07QP1&cluster=8845.f. Contact :			
Peng Liang Email : fliang@lifetech.com URL :			
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600			
Faraday Avenue Genoscope sequence ID : CS0DJ010CG07QP1.			
Location/Qualifiers			
1..1022			
source			


```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS03101YN13"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

```

ORIGIN

```

Query Match      38.9%; Score 586; DB 9; Length 1022;
Best Local Similarity 100.0%; Pred. No. 1.6e-93;
Matches 586; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 920 CCCACCCCTTATGGCTTGCCTCTGCCTCTGGATTTTAGGTGTGATTTCTGGATCAGC 979
Db 367 CCCACCCCTTATGGCTTGCCTCTGCCTCTGGATTTTAGGTGTGATTTCTGGATCAGC 426
QY 980 GGATACCACTTCTCTCCACACCCGACGAGGGCTAGAAAATTTGTTGAGATTTTAT 1039
Db 427 GGATACCACTTCTCTCTCCACACCCGACGAGGGCTAGAAAATTTGTTGAGATTTTAT 486
QY 1040 ATCATCTTTGCAAAATGCTTCAGTTGTAAATGTGAAAATGGCTGGGAAAGAGGTGG 1099
Db 487 ATCATCTTTGCAAAATGCTTCAGTTGTAAATGTGAAAATGGCTGGGAAAGAGGTGG 546
QY 1100 TGTCCCTAATTTGTTTACTTGTAACTTGTCTTGTGCCCCCTGGGCACTTGGCCCTTTGTC 1159
Db 547 TGTCCCTAATTTGTTTACTTGTAACTTGTCTTGTGCCCCCTGGGCACTTGGCCCTTTGTC 606
QY 1160 TGCTCTCAGTGTCTTCCCTTGACATGGAAGAGGTGTGGCCAAATATCCCATCTTCT 1219
Db 607 TGCTCTCAGTGTCTTCCCTTGACATGGAAGAGGTGTGGCCAAATATCCCATCTTCT 666
QY 1220 TGCACCTCAAGCTGTGTGGCTCAGGGCTGGGGTGGCAGAGGAGGCGCTTACCTTATATC 1279
Db 667 TGCACCTCAAGCTGTGTGGCTCAGGGCTGGGGTGGCAGAGGAGGCGCTTACCTTATATC 726
QY 1280 TGTGTTGTTATCCAGGGCTCCAGACTTCCCTCTCTGCTGCCCCCACTGCACCCCTCTCCCC 1339
Db 727 TGTGTTGTTATCCAGGGCTCCAGACTTCCCTCTCTGCTGCCCCCACTGCACCCCTCTCCCC 786
QY 1340 CTTATCTATCTCTCTCTCGGCTCCCGAGCCAGCTTGGCTTCTTGTGCTCCCTCTCTGGGT 1399
Db 787 CTTATCTATCTCTCTCTCGGCTCCCGAGCCAGCTTGGCTTCTTGTGCTCCCTCTCTGGGT 846
QY 1400 CATCCCTCCACTGACTCTGACTATGGCAGCAGAACACACAGGGCCCTGGCCCTGGGATT 1459
Db 847 CATCCCTCCACTGACTCTGACTATGGCAGCAGAACACACAGGGCCCTGGCCCTGGGATT 906
QY 1460 TCATGTTGATCATTAATAAAGAAATTCGCAACCAAAAAA 1505
Db 907 TCATGTTGATCATTAATAAAGAAATTCGCAACCAAAAAA 952

```

RESULT 28

```

BM737591
LOCUS BM737591
DEFINITION K-EST0000094 S1SNU5 Homo sapiens cDNA clone S1SNU5-21-B02 5', mRNA
sequence.
ACCESSION BM737591
VERSION BM737591.1 GI:19058920
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 582);
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,

```

Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.
 21C Frontier Korean EST Project 2001
 Unpublished (2002)
 Contact: Kim YS
 Genome Research Center

Korea Research Institute of Bioscience & Biotechnology
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 21 row: B column: 02
 High quality sequence stop: 582.
 Location/Qualifiers

FEATURES
source

```

1..582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S1SNU5-21-B02"
/sex="F"
/tissue_type="Ascites"
/cell_type="Lymphoblast-like"
/lab_hosts="Top10P"
/clone_lib="S1SNU5"
/notes="Organ: Stomach; Vector: pCNS; Site 1: EcoRI; Site 2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tabacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoR I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10P by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

```

ORIGIN

```

Query Match      38.6%; Score 580.4; DB 12; Length 582;
Best Local Similarity 99.8%; Pred. No. 1.6e-92;
Matches 581; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 365 GTTCTCAGCACAGCGGTGTAGACATCTGCATCACATGTGGCTTGGCCCGGCTGACA 424
Db 1 GTTCTCAGCACAGCGGTGTAGACATCTGCATCAACAATGTGGCTTGGCCCGGCTGACA 60
QY 425 CCTGTCTCTCAGGCAGCACCCAGTGGTGTGAAGGACATGTTCAATGTCAACGTGTGCCCC 484
Db 61 CCTGTCTCTCAGGCAGCACCCAGTGGTGTGAAGGACATGTTCAATGTGAACGTGTGCCCC 120
QY 485 TCAGCATCTCCACACGGGAAGCCTACAGTCCATGAAGGAGCGGAATGTGGAGATGGGC 544
Db 121 TCAGCATCTCCACACGGGAAGCCTACAGTCCATGAAGGAGCGGAATGTGGAGATGGGC 180
QY 545 ACATCATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGTGTGACCCACT 604
Db 181 ACATCATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGTGTGACCCACT 240
QY 605 TCTATAGTCCACCAAGTATGCCGTGCTGACAGGGGACTGAGGAGAGCTTC 664
Db 241 TCTATAGTCCACCAAGTATGCCGTGCTGACAGGGGACTGAGGAGAGCTTC 300
QY 665 GGGAGGCCAGAGCCCAACATCCGAGCCACTGTGATCTCTCCAGGTGTGGTGGAGACAAAT 724
Db 301 GGGAGGCCAGAGCCCAACATCCGAGCCACTGTGATCTCTCCAGGTGTGGTGGAGACAAAT 360
QY 725 TCGCCTTCAAACTCCACGACAGAGGACCCTGAGAGGACGTGCCACCTATGAGCAATGA 784

```

Db	361	TCGCCTTCAAACTCCACGACAGACCCCTGAGAAGGAGCTGCCACCTATGAGCAATGA	420
Qy	785	AGTGTCTCAAACCCGAGGATGTGGCCGAGGCTGTTATCTACGTCTCAGCACCCCGCCAC	844
Db	421	AGTGTCTCAAACCCGAGGATGTGGCCGAGGCTGTTATCTACGTCTCAGCACCCCGCCAC	480
Qy	845	ACATCCAGATGGAGACATCCAGATGAGGCCCGACGAGGAGTACCTAGTACTGTGGG	904
Db	481	ACATCCAGATGGAGACATCCAGATGAGGCCCGACGAGGAGTACCTAGTACTGTGGG	540
Qy	905	AGCTCCTCCTCCCTCCCGCCCTTCATGGCTTCCTCCCTGC	946
Db	541	AGCTCCTCCTCCCTCCCGCCCTTCATGGCTTCCTCCCTGC	582
RESULT 29			
LOCUS	BE386240	601273447F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3614770 5',	
DEFINITION		mRNA linear EST 21-JUL-2000	
ACCESSION		mRNA sequence.	
VERSION	BE386240		
KEYWORDS	BE386240.1	GI:9331605	
SOURCE		EST.	
ORGANISM		Homo sapiens (human)	
REFERENCE		Homo sapiens	
AUTHORS		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
TITLE		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
JOURNAL		NIH-MGC http://mgc.nci.nih.gov/ .	
COMMENT		1 (bases 1 to 769)	
		National Institutes of Health, Mammalian Gene Collection (MGC)	
		Unpublished (1999)	
		Contact: Robert Strausberg, Ph.D.	
		Email: cgapbs-remail.nih.gov	
		Tissue Procurement: ATCC/DCTD/DTF	
		cDNA Library Preparation: Ling Hong/Rubin Laboratory	
		DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)	
		DNA Sequencing by: Incyte Genomics, Inc.	
		Clone distribution: MGC clone distribution information can be	
		found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov	
		Plate: LCM277 row: m column: 11	
		High quality sequence stop: 652.	
FEATURES		Location/Qualifiers	
source		1..769	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:3614770"	
		/tissue_type="melanotic melanoma"	
		/lab_host="DH10B (phage-resistant)"	
		/clone_lib="NIH_MGC_20"	
		/note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:	
		EcoRI; cDNA made by oligo-dT priming. Directionally	
		cloned into EcoRI/XhoI sites using the following 5'	
		adapter: GGACAGG(G). Size-selected >500bp for average	
		insert size 1.8kb. Library constructed by Ling Hong in	
		the laboratory of Gerald M. Rubin (University of	
		California, Berkeley) using ZAP-cDNA synthesis kit	
		(Stratagene) and Superscript II RT (Life Technologies)."	
ORIGIN			
Query Match		38.4%; Score 578.4; DB 10; Length 769;	
Best Local Similarity		94.6%; Pred. No. 3.6e-92;	
Matches 643; Conservative		0; Mismatches 31; Indels 6; Gaps 4;	
Qy	800	AGGATGTGGCCGAGGCTGTTATCTACGTCTCAGCACCCCGCCACATCCAGATTGGAG	859
Db	1	AGGATGTGGCCGAGGCTGTTATCTACGTCTCAGCACCCCGCCACATCCAGATTGGAG	60
Qy	860	ACATCCAGATGAGGCCCGACGAGCAGTACCTAGTACTGTGGAGCTCCTCCCTCCCT	919
Db	61	ACATCCAGATGAGGCCCGACGAGCAGTACCTAGTACTGTGGAGCTCCTCCCTCCCT	120
Qy	920	CCCCACCCCTTCATGGCTTGCCCTCCTGCCTCTGGATTTTAGGTGTTGATTTCTGGATCAGC	979

```

/sex="F"
/tissue_type="Ascites"
/cell_type="Lymphoblast-like"
/lab_line="SNU-5"
/lab_host="Top10F"
/clone_lib="SISNUSel"
/notes="Organ: Stomach; Vector: pCNS; Site 1: EcoRI;
Site 2: NotI; The poly (A) + RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dT-selected mRNA by
priming with dT-tailed vector. The dT-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library. After analyzing and
sequencing about 2,000 ~ 3,000 colonies in original cDNA
library, the abundant cDNAs were selected and amplified by
PCR reaction using vector region primer including T7
promotor as 5' primer and N(dT)14 as 3' primer. The PCR
products were used as template for synthesis of
biotinylated single stranded RNA by in vitro transcription
reaction. The synthesized RNA probes were hybridized with
antisense single stranded cDNAs prepared from original
library and incubated with avidin-gel. After removing
DNA-RNA hybrids by centrifuge, the substracted cDNA
libraries were constructed by transformation of the
remaining DNA into competent cells E. coli Top10F with
electroporation method."

```

ORIGIN

```

Query Match      38.1%; Score 573; DB 12; Length 573;
Best Local Similarity 100.0%; Pred. No. 3.3e-91;
Matches 573; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 350 TGTTCAGCTATCGTTCTCAGACACGCGGTGTAGACATCTGCATCAACAATCTGGCT 409
DB 1 TGTTCAGCTATCGTTCTCAGACACGCGGTGTAGACATCTGCATCAACAATCTGGCT 60
QY 410 TGGCCCGGCTGACACCTGCTCTCAGCAGCAGCAGTGGTTGAAGACATGTTCAATG 469
DB 61 TGGCCCGGCTGACACCTGCTCTCAGCAGCAGCAGTGGTTGAAGACATGTTCAATG 120
QY 470 TGAACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCCTTACCAGTCCATGAAGGAGCGGA 529
DB 121 TGAACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCCTTACCAGTCCATGAAGGAGCGGA 180
QY 530 ATGTGACGATGGGCATCATTTAACTAATAGCATGTCTGGCCACCGAGTGTATCCCC 589
DB 181 ATGTGACGATGGGCATCATTTAACTAATAGCATGTCTGGCCACCGAGTGTATCCCC 240
QY 590 TGTCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTCATCTGGCTGACAGAGGAC 649
DB 241 TGTCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTCATCTGGCTGACAGAGGAC 300
QY 650 TGAGGCAAGAGCTTCGGGAGGCGCCAGACCCACATCCGAGCCAGTGTGATCTCTCCAGGTG 709
DB 301 TGAGGCAAGAGCTTCGGGAGGCGCCAGACCCACATCCGAGCCAGTGTGATCTCTCCAGGTG 360
QY 710 TGTGTGAGACACAATTCGCTTCTCAAACTCCAGCAAGGACCTTGAGAGGCGAGCTGCCA 769
DB 361 TGTGTGAGACACAATTCGCTTCTCAAACTCCAGCAAGGACCTTGAGAGGCGAGCTGCCA 420
QY 770 CCTATGAGCAAAATGAATGTCTCAAAACCCAGAGATGTGGCCGAGGCTGTATCTACGTCC 829
DB 421 CCTATGAGCAAAATGAATGTCTCAAAACCCAGAGATGTGGCCGAGGCTGTATCTACGTCC 480

```

```

QY 830 TCAGCACCCCCCCACACATTCAGATTGGAGACATCCAGATGAGGCCACCGAGCAGGTGA 889
DB 481 TCAGCACCCCCCCACACATTCAGATTGGAGACATCCAGATGAGGCCACCGAGCAGGTGA 540
QY 890 CCTACTGACTGTGGGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 922
DB 541 CCTACTGACTGTGGGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 573

RESULT 31
BE869917
LOCUS BE869917
DEFINITION 60144563F1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3850574 5',
mRNA sequence.
ACCESSION BE869917
VERSION BE869917.1 GI:10318693
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 614)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9569 row: n column: 15
High quality sequence stop: 602.
FEATURES
Location/Qualifiers
1..614
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3850574"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 65"
/notes="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.8 kb. Library constructed by Life
Technologies."

```

ORIGIN

```

Query Match      37.7%; Score 567.4; DB 10; Length 614;
Best Local Similarity 99.2%; Pred. No. 3.2e-90;
Matches 612; Conservative 0; Mismatches 1; Indels 4; Gaps 4;

QY 33 GGACGAGAGCGCCCGGGCTCAGCTCTCTGACCCCGGCTGTCTCGGGTAGTCCAGCGAGGC 92
DB 1 GGACGAGAGCGCCCGGGCTCAGCTCTCTGACCCCGGCTGTCTCGGGTAGTCCAGCGAGGC 60
QY 93 GGACGCGGCGCTGGGCGCCATGCGCCAGGCCCGGCGATGGCGCGGTGGCGGACCGGTGGC 152
DB 61 GGACGCGGCGCTGGGCGCCATGCGCCAGGCCCGGCGATGGCGCGGTGGCGGACCGGTGGC 120
QY 153 GTTGTGTGACGGGGGCTCTCGGGGGGCGATCGCGCGGCGCTGGCCCGGCGCTGTCCAGCA 212
DB 121 GTTGTGTGACGGGGGCTCTCGGGGGGCGATCGCGCGGCGCTGGCGCGGCGCTGTCCAGCA 179
QY 213 GGGACTGAAGGTGGTGGGCTGCGCCCGCACTGTGGGCAACATCGAGGAGCTGGCTCTCA 272
DB 180 GGGACTGAAGGTGGTGGGCTGCGCCCGCACTGTGGGCAACATCGAGGAGCTGGCTCTCA 239
QY 273 ATGTAAAGTGTGAGGCTACCCCGGAGCTTTGTATCCCTACAGATGTGACCTATCAATGA 332

```

priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells *E. coli* Top10F⁺ by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library. After analyzing and sequencing about 2,000 ~ 3,000 colonies in original cDNA library, the abundant cDNAs were selected and amplified by PCR reaction using vector region primer including T7 promoter as 5' primer and Nid114 as 3' primer. The PCR products were used as template for synthesis of biotinylated single stranded RNA by *in vitro* transcription reaction. The synthesized RNA probes were hybridized with antisense single stranded cDNAs prepared from original library and incubated with avidin-gel. After removing cDNA-RNA hybrids by centrifuge, the subtracted cDNA libraries were constructed by transformation of the remaining DNA into competent cells *E. coli* Top10F⁺ with electroporation method."

Query Match 37.6%: Score 566: DB 12: Length 566;

```
Query Match      37.6%; Score 566; DB 12; Length 566;
Best Local Similarity 100.0%; Pred. No. 5.7e-90;
Matches 566; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY
350 TGTTCACGATATCCGTTCTCAGCACAGCGGTGTAGACATCTGCATCAACAATGCTGGCT 409

Dd
1 TGTTCACGATATCCGTTCTCAGCACAGCGGTGTAGACATCTGCATCAACAATGCTGGCT 60

QY 410 TGGCCCGGCGCTGACACCCCTGCTCTCAGGAGCACCAGTGGTTGGAAGGACATGTTCAATG 469

pb 61 TGGCCCGGCGCTGACACCCCTGCTCTCAGGAGCACCAGTGGTTGGAAGGACATGTTCAATG 120

[illegible]

QY 470 TGAACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCCTACCCAGTCCATGAAAGGAGCGGA 529

180

Db 121 TGAACGTGCTGGCCCTCAGCATCTGCACACGGGAHGCCCTACCAAGTCCTCAAGGAGGAGGCTG

520 ATCTCGACGATCGGCACATCATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCC 589

[illegible]

Db
181 ATGTGGACGATGGGCACATCATTAACATCAATAGCATGCTGGCCACCGAGTGTACCCC 240

QY 590 TGTCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTCACTGCCGTGACAGAGGGAC 649

300

D6
241 TGTCTGTGACCCACTTCTATAGTGCCACCACCAAGTAATGCCGCACAGCGCTGACAGAGGGGAC

650 TGAGGCAAGAGCTTCGGGAGGCCCCAGACCCACATCCGAGCCACGTGCATCTCTCCAGGTG 709

[illegible]

Db 301 TGAGGCAAGAGCTTCGGGAGGCCAGACCCACATCCGAGCCACGTGCATCTCTCCAGGTG 360

[illegible]

QY 710 TGGTGGAGACACAAATTCGCCCTTCAAACTCCACGACAAGGACCTTGAGAAAGGCAGCTGCCA 769

TCCTGAGAGGAGCTGCA 420

DB TGGTGGAGACACAAATTCGGCTTCATATCCACGACAGGAGACCCATGAGATAGCCAGCCAGCCAT

770 CCTATGAGCAAAATGAAGTGTCTCAAAACCCGAGGATGTGGCCGAGGCTGTATCTACGTCC 829

Db 421 CCTATGAGCAAATGAAGTGTCTCAAACCGAGGATGTGGCCGAGGCTGTATCTACGTCC 480

QY
830 TCAGCACCCCGCACACATCCAGATTGGAGACATCCAGATGAGGCCACGGAGCAGTGA 889

TCTCCGCGCAGTCACATCCGCTCAGGCCACGGAGCAGGTGA 540

DB 481 T CAGCACCCCGGACACATCCAGATGGAGACATCCAGATGAAAGGCCAGCGGACGCTGT

890 CCTAGTGACTGTGGGAGCTCCTCCTT 915

[illegible]

Db 541 CCTAGTGA CTGTGGGAGCTCCTCCTT 566

RESULT 33
=====

BM682228/c					
TGCTC	PMCGCCCG	572 bp	mRNA	linear	EST 27-FEB-200

LOCUS
BM622228
UI-E-EJ0-aio-j-09-0-UI.s1 UI-E-EJ0 Homo sapiens cDNA clone
DEFINITION

2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841

UI-E-EJ0-ai0-j-09-0-UI 3', mRNA sequence.
 BM682228
 VERSION
 BM682228.1 GI:18992124
 EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 572)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 The following repetitive elements were found in this cDNA
 sequence: 1-38, >POLY_A#Simple_repeat (matched complement)
 Seq primer: M13 Forward
 POLYA=Yes.

FEATURES
 source
 1..572
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_lib="UI-E-EJ0"
 /tissue="UI-E-EJ0-ai0-j-09-0-UI"
 /tissue_type="fetal eyes, lens, eye anterior segment,
 optic nerve, retina, Retina Foveal and Macular, RPE and
 Choroid"
 /dev_stage="fetal and adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-E-EJ0"
 /note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
 modified polylinker; Site_1: EcoR I; Site_2: Not I;
 UI-E-EJ0 is a subtracted cDNA library constructed
 according to Bonaldo, Lennon and Soares, Genome Research,
 6:791-806, 1996. First strand cDNA synthesis was primed
 with an oligo-dT primer containing a Not I site. Double
 stranded cDNA was ligated to an EcoR I adaptor, digested
 with Not I, and cloned directionally into pT73-Pac
 vector. The oligonucleotide used to prime the synthesis of
 first-strand cDNA contains a library tag sequence that is
 located between the Not I site and the (dT)18 tail. The
 sequence tags for this library are: fetal eyes,
 AGATCAAGA; lens, CGATTAGCGA; eye anterior segment,
 AATGCCGCAT; optic nerve, CCATTAAAGTG; retina, CCGCG;
 Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This
 library was created for the program, Gene Discovery in the
 Visual System, supported by National Eye Institute (NEI).
 TAG TISSUE=human retina
 TAG LIB=UI-E-EJ0
 TAG_SEQ=CCGCG"

ORIGIN
 Query Match 37.5%; Score 565; DB 12; Length 572;
 Best Local Similarity 99.8%; Pred. No. 8.6e-90;
 Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 940 CTCCTCGCTGGATTTAGGTTGATTTCTGGATCAGGGATACACTTCCTGTCAC 999
 DB 572 CTCCTCGCTGGATTTAGGTTGATTTCTGGATCAGGGATACACTTCCTGTCAC 513

QY 1000 ACCCGACGAGGGCTAGAAAAATTTGTTTGTAGATTTTATATCATCTTGTCAAATGCTT 1059
 DB 512 ACCCGACGAGGGCTAGAAAAATTTGTTTGTAGATTTTATATCATCTTGTCAAATGCTT 453
 QY 1060 CAGTTGTAATGTGAATAATGGCTGGGAAAGGAGGTGTCTCCCTAATGTTTACTT 1119
 DB 452 CAGTTGTAATGTGAATAATGGCTGGGAAAGGAGGTGTCTCCCTAATGTTTACTT 393
 QY 1120 GTTAACTTGTCTTGTGCCCCCTGGGCACCTTGGCCCTTGTCTCTCAGTGTCTTCCCTT 1179
 DB 392 GTTAACTTGTCTTGTGCCCCCTGGGCACCTTGGCCCTTGTCTCTCAGTGTCTTCCCTT 333
 QY 1180 TGACATGGGAAAGGAGTTGTGGCCAAAATCCCACTTCTTTCACACTCAACGCTGTGTGGC 1239
 DB 332 TGACATGGGAAAGGAGTTGTGGCCAAAATCCCACTTCTTTCACACTCAACGCTGTGTGGC 273
 QY 1240 TCAGGGCTGGGTTGTCAGAGGAGGAGGCTTCACTTATATCTGTGTGTATCCAGGGCTC 1299
 DB 272 TCAGGGCTGGGTTGTCAGAGGAGGAGGCTTCACTTATATCTGTGTGTATCCAGGGCTC 213
 QY 1300 CAGACTTCTCTCTGCTGCCCCCTGCACCTGCACCTCTCCCCCTTATCTATCTCTCTCGG 1359
 DB 212 CAGACTTCTCTCTGCTGCCCCCTGCACCTGCACCTCTCCCCCTTATCTATCTCTCTCGG 153
 QY 1360 CTCGCCAGCCAGTCTGGCTTGTGTCCTCTCTGGGGTTCATCCCTCCACTCTGACTCT 1419
 DB 152 CTCGCCAGCCAGTCTGGCTTGTGTCCTCTCTGGGGTTCATCCCTCCACTCTGACTCT 93
 QY 1420 GACTATGCGACAGAACACGAGGCGCTGGCCAGTGGATTTCATGGTGATCATTAATAAAA 1479
 DB 92 GACTATGCGACAGAACACGAGGCGCTGGCCAGTGGATTTCATGGTGATCATTAATAAAA 33
 QY 1480 GAAAAATCGCAACCAAAAAA 1505
 DB 32 GAAAAATCGCAACCAAAAAA 7

RESULT 34
 BI761329
 LOCUS BI761329 748 bp mRNA linear EST 25-SEP-2001
 DEFINITION 603044093F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5184479 5',
 mRNA sequence.
 ACCESSION BI761329
 VERSION BI761329.1 GI:15752907
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 748)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-i@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLAM1460 row: i column: 24
 High quality sequence stop: 724.
 Location/Qualifiers
 1..748
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5184479"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_116"

/notes="Organ: pooled colon, kidney, stomach; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV; Detected: source anonymous_pool of 3 colons, age 26 yo male, 49 yo female, 71 yo male colon; 46 yo female kidney, and pool of 2 stomachs, 62 yo male and 70 yo female. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 023. Note: this is a NIH MGC Library."

ORIGIN

Query Match	36.6%	Score 550.2	DB 12	Length 748
Best Local Similarity	93.9%	Pred. No. 3.5e-87		
Matches 630	Conservative 0	Mismatches 28	Indels 13	Gaps 5
Qy	10	GACCCAAAGCAGGTGCGGGCGCGCGCAGGAGAGCGCGCGGCGGTGAGTCTCTCGACACCCC	69	
Db	70	GACCCAAAGCAGGTGCGGGCGCGCGCAGGAGAGCGCGCGGTGAGTCTCTCGACACCCC	129	
Qy	70	GTGTCGGGCTAGTCCAGCGAGG-----CGGACGGGCGGCGTGGGCCCATGGCCAGGCC	122	
Db	130	GTGTCGGGCTAGTCCATCGAGGCGTGACTGGGTGGCGTGGTGCCATGTGSCAGGCC	189	
Qy	123	CGGCATGAGAGCGGTG---GCCGACCGCGTGGCGCTGGTGACGGGGGCGCTCGGGGGGCAT	179	
Db	190	TGGCATGGAGCGGTGTGCGCGACCTGTGCTGGCGTGGTGACGGGGGCGCTCGGGGGGCAT	249	
Qy	180	CGGCGCGGCGTGGCCCC--GGGCCCTGTGTCAGACGGGACTGAAGGTGGTGGGCTGCGCCC	238	
Db	250	CGGCGCGGCGTGGCCCCCTGGGCTGTCAGACGGGACTGAAGGTGGTGGGCTGCGCCC	309	
Qy	239	GCACGTGTGGGCAACATCGAGAGGTGGTGTCTGAATGTAAAGTGCAGGCTACCCCGGGA	298	
Db	310	GCACGTGTGGGCAACATCGAGAGGTGGTGTCTGAATGTAAAGTGCAGGCTACCCCGGGA	369	
Qy	299	CTTTGATCCCCCTACAGATGTGACCTATCAAATGAAGAGGACATCTCTCCATGTTTCTCAG	358	
Db	370	CTTTGATCCCCCTACAGATGTGACCTATCAAATGAAGAGGACATCTCTCCATGTTTCTCAG	429	
Qy	359	CTATCCGGTTCACGACAGCGGTGAGACATCTGCATCAACAATGCTGGCTGGGCCCGGC	418	
Db	430	CTATCCGGTTCACGACAGCGGTGAGACATCTGCATCAACAATGCTGGCTGGGCCCGGC	489	
Qy	419	CTGACACCCCTGCTCTCAGCGAGCACACAGTGGTTGGAAGGACATGTTTCAATGTGAACGTGC	478	
Db	490	CTGAACCCCTGCTCTCAGCGAGCACACAGTGGTTGGAAGGACATGTTTCAATGTGAACGTGC	549	
Qy	479	TGGCCCTCAGCATC--TGACACAGGGAAGCCTACAGTCCATGAAGAGCGGAATGTGGA-	536	
Db	550	TGGCCCTCAGCATCTTGACACAGGGAAGCCTACCAAGTCCATGAAGAGCGGAATGTGGA	609	
Qy	537	CGATGGGCACATCAATTAACTAAATGAGATGTCGCGCACCGAGTGTACCCCTGTCTGT	596	
Db	610	CGATGGGCACATCAATTAACTAAATGAGATGTCGCGCACCGAGTGTACCCCTGTCTGT	669	
Qy	597	GACCCACTTCTATAGTGGCCACCAAGTATGCGGCTCACTGCGCTGACAGAGGACTGAGGCA	656	
Db	670	GACCCACTTCTATAGTGGCCACCAAGTATGCGGCTCACTGCGCTGACAGAGGACTGAGGCA	729	
Qy	657	AGAGCTTCGGG	667	
Db	730	CGAGCTTCGGG	740	

RESULT 35
BG967186 914 bp mRNA linear EST 12-JUN-2001
LOCUS 602833867f1 NCI_CGAP-Co2a Mus musculus cDNA clone IMAGE:4986176 5,
DEFINITION mRNA sequence.
ACCESSION BG967186
VERSION BG967186.1 GI:14354823

KEYWORDS	REFERENCES
SOURCE	AUTHORS
ORGANIS	TITLE
	JOURNAL
	COMMENT

EST.
Miss miscellanea (house mouse)

the macellus
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 914)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cdDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.

<http://image.llnl.gov>

plate: LLAM11000 row: 1 column: 1
with multi: commence stop: 748

High quality sequence scop: 140
 Location/Qualifiers

FEATURES

T-078

"Organism"-Mia miscellanea"

```
/organism="Mab"
/mol_type="mRNA"
```

```

/strain="FVB/N"

```

```
/db xref="taxon:10090"
```

/clone="IMAGE:4988176"

```
/lab host="DH10B (T1 phage-resistant)"
```

```
/clone lib="NCI CGAP Co24"
```

```
/note="Organ: colon; Vector: pCMV-SPORT"
```

Site 2: SalI; Cloned unidirectionally.

Average insert size 1.6 kb. Constructed

Technologies. Note: this is a NCI_CGAF

ORIGIN

Query Match	36.2%	Score	544.2	DB	12	Length	914
Best Local Similarity	85.4%	Pred.	No. 4e-86				
Matches	734	Conservative	0	Mismatches	113	Indels	12
Gaps	11						
QY	66	CCCCGTGTCGGGCTAGTCCACGAGGCGGGA	CGGCGCGTGGGGCCCATGGCCAGGCCCGG	125			
DB	59	CCCAGTGCCTGGTTCGCCGACGAGGTGGG	CAGACGGCTCGGACCAATGACTAGAGCTGG	118			
QY	126	CATGAGCGGTGGGCGGACACCGGCTGGG	CGCTGGTGACGGGGGCGCTCGGGGGGCATCGGCGC	185			
DB	119	CATGAGCGGTGGGCGGACACCGGCTGGCA	CTGGTACGCGGAGCCTCGGGGGGCATCGGTGC	178			
QY	186	-GGCGTGGCCCGGCGCCCTGGTCAGACGG	GACCTGAAGTGGTGGGCTCGGCCGCCCACTG	244			
DB	179	TGGCGGTGGCCCGGCGCATTAGTCACAGC	AGGACTGAAGTGTGGGTGTGGGTGTGCCGCCACCG	238			
QY	245	TGGGCAACATCGAGGAGCTGGCTGCTGAA	TGAA-GAGTCAGGCTACCCCGGGACTTTG	303			
DB	239	TTGCAACATCGAGGAGCTGGCTGCTGAA	TGTAACTGACGAGTCAGGCTACCCCGGGACTTTG	298			
QY	304	ATCCCTACAGATGTGACCTATCAATGAAG	GAGNACCTCTCCATGTTCTCAGCTATC	363			
DB	299	ATCCCTACAGATGTGACCTGTCAATGAGG	AGGACATCCCTCTCCATGTTCTCAGCTGTC	358			
QY	364	CGTTCTCAGCACAGCGGTGTAGACTCTGC	ATCTGCATCAACAATGCTGGCTGGCCCGGCTGCAC	423			
DB	359	CGATCCCAGCACAGTGGGTGGATATCTGC	ATCATCAACAATGCGGGATGCGCCGGCTGCAC	418			
QY	424	ACCCTGCTCTCAGGCGAGCACAGTCGGT	TGGAAGGACATGTTCAATGTGAACGTGCTGGCC	483			
DB	419	ACCCTGCTCTCGGGCAGCACACCGGATG	GGAAGGACATGTTCAATGTGAATGTGCTGGCC	478			
QY	484	CTCAGCATCTGACACGCGGAGCCTACCA	GTCCTACCGAGCGGAATGTGCACGATGG	542			
DB	479	CTCAGCATCTGACCTCGGAGGCTTATCAG	TCCATCTGACGAGCGGAACATAGACGACGG	538			
QY	543	GCACATCATTAACATCAATAGCATGTCT	TGGCCACCGGAGTTTACCCCTCTCTGTGTGACCCA	602			


```

Db      539 GCACATCAATTAACATCAACAGCATGTGTGGCCACCGAGTCCACCCAGTCTGTGATCCA 598
QY      603 CTTCTATAGTCCACCAAGTATCCCTCACTGCGCTGACAGAGGACTGAGGCAAGACT 662
Db      599 TTCTATAGTGGACTAAGTATGCCCTCACTGCACCTGACAGAGGACTCAGGC-AGAGCT 657
QY      663 TCGGAGGGCCAGACCCACCATCCAGCCACGTGTCATCTCTCCAGGTGTG-GTGAGACAC 721
Db      658 TCTGGAGGGCCAGACCCACCATATCCG-GGCACGTGTATCTCTCCAGGTTCGGTAGAGAC 716
QY      722 AATTGCGCTTCAACTCCAGCACAGGACCC-TGAGAAGGCGAGCTG-CCACTATGAGCA 779
Db      717 AGTTGCGCTTCAACTCCAGCACAGGACCCGGGCGAAGCGAGTCTGACGTGCTTAGCACA 776
QY      780 AAT-GAAGTGTCTCAAAACCCGAGGAGTGTGGCGAGGCTGTAT-CTACGTCTCTCAGCAC 837
Db      777 CTTACAGTGTCTCAGACCAGAGGAGTGTCTGAGGCTGTATACGTGCTTAGCACA 836
QY      838 CCGCACACATCCAGATTGGAGACATCCAGATGAGGCCACCGAGCAGGTGACCTAGTGA 897
Db      837 ACCCCACATGTGAGGCTTGGGACATCCAGATGAGGCCACAGACGAG--TGACTTATGT 894
QY      898 CTGTGGGAGCTCTCTCTTC 916
Db      895 CTGTGGGAGATCCTACTCC 913

RESULT 36
BM675954/c
LOCUS
DEFINITION
  UI-E-EJ0-ahf-o-21-0-UI.s1 UI-E-EJ0 Homo sapiens cDNA clone
  UI-E-EJ0-ahf-o-21-0-UI 3', mRNA sequence.
ACCESSION
  BM675954
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 570)
  Bonaldo,M.F., Lennon,G. and Soares,M.B.
  Normalization and subtraction: two approaches to facilitate gene
  discovery
  Genome Res. 6 (9), 791-806 (1996)
  97044477
  8889548
  Contact: Soares, MB
  Coordinated Laboratory for Computational Genomics
  University of Iowa
  375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
  Tel: 319 335 8250
  Fax: 319 335 9565
  Email: bento-soares@uiowa.edu
  Tissue Procurement: Dr. Gregg Hageman
  cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
  cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
  DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
  Clone Distribution: Researchers may obtain clones from Research
  Genetics (www.resgen.com).
  The following repetitive elements were found in this cDNA
  sequence: 1-38, >POLY_A#Simple_repeat (matched complement)
  Seq primer: M13 Forward
  POLYA=Yes.
FEATURES
  Location/Qualifiers
    1..570
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="UI-E-EJ0-ahf-o-21-0-UI"
      /tissue_type="fetal eyes, lens, eye anterior segment,
      optic nerve, retina, Retina Foveal and Macular, RPE and
      Choroid"
      /dev_stage="fetal and adult"

```

```

/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-EJ0"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-EJ0 is a subcloned cDNA library constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT7T3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tags for this library are: fetal eyes,
AATGCCGAT; lens, CGATTAGCGA; eye anterior segment,
AATGCCGAT; optic nerve, CCATTAGTGC; retina, CCGCG; Retina
Foveal and Macular, GTCC; RPE and Choroid, ACTTA. This
library was created for the program, Gene Discovery in the
Visual System, supported by National Eye Institute (NEI).
TAG TISSUE=human retina
TAG LIB=UI-E-EJ0
TAG_SEQ=CCGCG"

```

ORIGIN

```

Query Match      36.1%; Score 543; DB 12; Length 570;
Best Local Similarity 99.8%; Pred. No. 6.7e-86;
Matches 554; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY      951 GGATTTAGGTGTTGATTCTCTGGATCAGGATACCACTTCTCTGCACACCCGACAG 1010
Db      560 GGATTTAGGTGTTGATTCTCTGGATCAGGATACCACTTCTCTGCACACCCGACAG 501
QY      1011 GGGCTAGAAAAATTTGTTTGAGATTTTATATCATCTTGTCAAATTTGCTTCAAGTTGTAAT 1070
Db      500 GGGCTAGAAAAATTTGTTTGAGATTTTATATCATCTTGTCAAATTTGCTTCAAGTTGTAAT 441
QY      1071 GTGAAAAATGGCTGGGAAAGAGGTGGTGCCTAATGTTTACTTGTAACTTGT 1130
Db      440 GTGAAAAATGGCTGGGAAAGAGGTGGTGCCTAATGTTTACTTGTAACTTGT 381
QY      1131 CTTCTGCCCCCTGGGCACCTTGGCCCTTCTGCTCTCAAGTCTTCCCTTTGACATGGGAA 1190
Db      380 CTTCTGCCCCCTGGGCACCTTGGCCCTTCTGCTCTCAAGTCTTCCCTTTGACATGGGAA 321
QY      1191 AGGAGTTGTGGCCAAAAATCCCACTCTTTTGACCTCAAGCTCTGTGGCTCAGGCTGGG 1250
Db      320 AGGAGTTGTGGCCAAAAATCCCACTCTTTTGACCTCAAGCTCTGTGGCTCAGGCTGGG 261
QY      1251 GTGGCAGAGGAGGCGCTTACCTTATATCTGTGTGTTTATCCAGGGCTCCAGACTTCTC 1310
Db      260 GTGGCAGAGGAGGCGCTTACCTTATATCTGTGTGTTTATCCAGGGCTCCAGACTTCTC 201
QY      1311 CTCTGCTGCCCCCAGCTGCACCTCTCCCTTATCTATCTCTCTCTGCTCCGCTCCCGAGCC 1370
Db      200 CTCTGCTGCCCCCAGCTGCACCTCTCCCTTATCTATCTCTCTCTGCTCCGCTCCCGAGCC 141
QY      1371 AGCTTTGGCTTCTTGTGCCCCCTCTCTGGGGTCACTCCCTCCACTCTGACTCTGACATGGCAG 1430
Db      140 AGCTTTGGCTTCTTGTGCCCCCTCTCTGGGGTCACTCCCTCCACTCTGACTCT-ACTATGGCAG 82
QY      1431 CAGAACACAGGCGCTGGCCCGAGTTCATGTTTCATGGTGATCATTAATAAAAAAATCGCA 1490
Db      81 CAGAACACAGGCGCTGGCCCGAGTTCATGTTTCATGGTGATCATTAATAAAAAAATCGCA 22
QY      1491 ACCAAAAAATAAAAA 1505
Db      21 ACCAAAAAATAAAAA 7

```

RESULT 37
 BG231973/c
 LOCUS
 DEFINITION

BG231973 541 bp mRNA linear EST 09-FEB-2001
 naf34g12.x1 Soares_NPBMc Homo sapiens cDNA clone IMAGE:4143166 3',

QY	mRNA sequence.		1312	TCCTGCTGCCCACTGCACCTCTCCCTCTATCTATCTCTCTCGCTCCCGCCCA	1371
	ACCESSION	BG231973			
	VERSION	BG231973.1	GI:12727118		
	KEYWORDS	EST.	181	TCCTGCTGCCCACTGCACCTCTCCCTCTATCTATCTCTCTCGCTCCCGCCCA	122
	SOURCE	Homo sapiens (human)			
QY	Homo sapiens		1372	GTCTTGGCTTTCTTGTCTCCCTCTCCCTGCGGTCTATCCCTCCACTCTGACTCTGACTATGGCAGC	1431
	ORGANISM	Homo sapiens			
	REFERENCE	1 (bases 1 to 541)			
	AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .			
	TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),			
QY	Tumor Gene Index		1432	AGAACACCGGCTGGCCAGTGGATTTCATGTGATCATTTAAAGAAAGAAATCGCAA	1491
	JOURNAL	Unpublished (1997)			
	COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-re@mail.nih.gov Tissue Procurement: M. Bento Soares, Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D. and M. Fatima Bonaldo, Ph.D. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov			
		Seq primer: -400P from Gibco			
		High quality sequence stop: 480.			
QY	Location/Qualifiers		61	AGAACACCGGCTGGCCAGTGGATTTCATGTGATCATTTAAAGAAAGAAATCGCAA	2
	FEATURES	source	1492	C 1492	
		1. 541			
		/organism="Homo sapiens"			
		/mol_type="mRNA"			
QY	db_xref="taxon:9606"				
		/clone="IMAGE:414316"			
		/tissue_type="lymphocyte"			
		/lab_host="DH10B (phage-resistant)"			
		/clone_lib="Soares NPBC"			
QY	/note="Organ: blood; Vector: pTV73D-Pac; Site 1: NotI; Site 2: EcoRI; 1st strand cDNA was primed with a Not I - oligo (dT) primer [5', TGTTACCAATCTGAAGTGGAGCGGCCCGGTTTTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTV73 vector. Library is normalized; constructed in the laboratory of M. Bento Soares (University of Iowa)."				
	ORIGIN				
	Query Match	35.9%; Score 541; DB 12; Length 541;			
	Best Local Similarity	100.0%; Pred. No. 1.5e-85;			
	Matches	541; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	952 GATTTAGGTGTGATTTCTGGATCAGGGATACCACTTCCTGTCACACCCGACCGG				1011
QY	541 GATTTAGGTGTGATTTCTGGATCAGGGATACCACTTCCTGTCACACCCGACCGG				482
QY	1012 GCTAGAAATTTGTTGAGATTTTATATCATCTTGTCAAATGCTTCAGTTGTAATG				1071
QY	481 GCTAGAAATTTGTTGAGATTTTATATCATCTTGTCAAATGCTTCAGTTGTAATG				422
QY	1072 TGAAAAATCGGCTGGGGAAGAGGTGGTGCCTAATTTGTTTACTGTTAATCTGTTTC				1131
QY	421 TGAAAAATCGGCTGGGGAAGAGGTGGTGCCTAATTTGTTTACTGTTAATCTGTTTC				362
QY	1132 TTGTGCCCCGACCTTGGCCCTTGTCTGCTCTCAGTCTTCCTTTGACATGGGAA				1191
QY	361 TTGTGCCCCGACCTTGGCCCTTGTCTGCTCTCAGTCTTCCTTTGACATGGGAA				302
QY	1192 GGAGTTGTCGCAAAATCCCCTCTCTTGCACTCAACGCTCTGTGGCTCAGGCTGGGG				1251
QY	301 GGAGTTGTCGCAAAATCCCCTCTCTTGCACTCAACGCTCTGTGGCTCAGGCTGGGG				242
QY	1252 TGGCAGAGGAGGCGCTTCACCTTATATCTGTGTGTATCCAGGCTCCAGACTTCTCC				1311
QY	241 TGGCAGAGGAGGCGCTTCACCTTATATCTGTGTGTATCCAGGCTCCAGACTTCTCC				182

library, the abundant cDNAs were selected and amplified by PCR reaction using vector region primer including T7 promoter as 5' primer and N(ΔT)14 as 3' primer; the PCR products were used as template for synthesis of biotinylated single stranded RNA by in vitro transcription reaction. The synthesized RNA probes were hybridized with antisense single stranded cDNAs prepared from original library and incubated with avidin-gel. After removing DNA-RNA hybrids by centrifuge, the subtracted cDNA libraries were constructed by transformation of the remaining DNA into competent cells *E. coli* Top10⁺ with electroporation method."

ORIGIN

```

Query Match      35.8%; Score 539; DB 12; Length 550;
Best Local Similarity 99.8%; Pred. No. 3.4e-85;
Matches 550; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 350 TGTTCACGCTATCGCTTCTCAGCAGCGGTGTAGACATCTGCATCAACAATGCTGGCT 409
    |||||
Db 1 TGTTCACGCTATCGCTTCTCAGCAGCGGTGTAGACA-CTGCATCAACAATGCTGGCT 59

QY 410 TGGCCCGGCTGACACCTGCTCTCAGCAGCAGCAGTGTGGAGGACATGTTCAATG 469
    |||||
Db 60 TGGCCCGGCTGACACCTGCTCTCAGCAGCAGCAGTGTGGAGGACATGTTCAATG 119

QY 470 TGAACGCTGTGGCCCTCAGCATCTGCACACGGAGCCTACCAGTCCATGAGGAGCGGA 529
    |||||
Db 120 TGAACGCTGTGGCCCTCAGCATCTGCACACGGAGCCTACCAGTCCATGAGGAGCGGA 179

QY 530 ATGTGGAGCATGGGCATCATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCC 589
    |||||
Db 180 ATGTGGAGCATGGGCATCATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCC 239

QY 590 TGTCTGTACCCACTTCTATAGTGCCACCAAGTATGCCGTACTCTGGCTGACAGAGGAC 649
    |||||
Db 240 TGTCTGTACCCACTTCTATAGTGCCACCAAGTATGCCGTACTCTGGCTGACAGAGGAC 299

QY 650 TGAGCAGAGCTTGGGAGGCGCCAGACCCACATCCGAGCCACGCGCATCTCTCCAGTG 709
    |||||
Db 300 TGAGCAGAGCTTGGGAGGCGCCAGACCCACATCCGAGCCACGCGCATCTCTCCAGTG 359

QY 710 TGGTGGAGACAAATTCGCTTCAAACTCCAGCAAGGACCTTGAGAGGCGAGCTGCCA 769
    |||||
Db 360 TGGTGGAGACAAATTCGCTTCAAACTCCAGCAAGGACCTTGAGAGGCGAGCTGCCA 419

QY 770 CCTATGAGCAAAATGAAGTGTCTCAACCCGAGGATGTGGCGAGGCTGTTATCTACGTCC 829
    |||||
Db 420 CCTATGAGCAAAATGAAGTGTCTCAACCCGAGGATGTGGCGAGGCTGTTATCTACGTCC 479

QY 830 TCAGCAGCCCGGCACATCCAGATTTGAGACATCCAGATGAGCCACCGAGCAGGTGA 889
    |||||
Db 480 TCAGCAGCCCGGCACATCCAGATTTGAGACATCCAGATGAGCCACCGAGCAGGTGA 539

QY 890 CCTAGTCACTG 900
    |||||
Db 540 CCTAGTCACTG 550

```

RESULT 39

```

CA424831/c
LOCUS
DEFINITION
  UI-H-FEI-bdw-e-17-0-UI.s1 NCI CGAP FEI Homo sapiens cDNA clone
  UI-H-FEI-bdw-e-17-0-UI 3', mRNA sequence.
ACCESSION
  CA424831
VERSION
  CA424831.1 GI:24787557
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 566)
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

```

TITLE

JOURNAL

COMMENT

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: csapbs@mail.nih.gov
Tissue Procurement: James Martin
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=tes.

FEATURES

source

```

1..566
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FEI-bdw-e-17-0-UI"
/tissue_type="Cell lines"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP FEI"
/notes="Organ: Chondrosarcoma; Vector: pT73-Pac
(Pharmacia) with a modified polylinker; Site 1: EcoR I;
Site 2: Not I; NCI CGAP FEI is a normalized cDNA library
derived from a pool of mRNA obtained from 3 cell lines
from grade II chondrosarcoma tissues. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dn)18 tail. The sequence tag for this library is
CGCTACGGAC. The cell lines were provided by Dr James
Martin from the University of Iowa.
TAG TISSUE=Human grade 2 chondrosarcoma cell line pool
TAG LIB=UI-H-FEI
TAG_SEQ=CGCTACGGAC"

```

ORIGIN

```

Query Match      35.4%; Score 533; DB 14; Length 566;
Best Local Similarity 98.6%; Pred. No. 3.9e-84;
Matches 548; Conservative 0; Mismatches 6; Indels 2; Gaps 1;

QY 952 GATTTTAGTGTGATTTCTGGATCAGGGATACCACTTCTGTCCACACCCGACGAG 1011
    |||||
Db 566 GATTTTAGTGTGATTTCTGGATCAGGGATACCACTTCTGTCCACACCCGACGAG 507

QY 1012 GGCTAGAAAAATTTGTTTGAGATTTTATATCATCTTGTCAATTCCTTCAGTTGTAATG 1071
    |||||
Db 506 GGCTAGAAAAATTTGTTTGAGATTTTATATCATCTTGTCAATTCCTTCAGTTGTAATG 447

QY 1072 TGAATAATGGCTGGGGAAGAGGAGTGGTCTCCCTAATTTGTTTACTTGTAACTTCCTC 1131
    |||||
Db 446 TGAATAATGGCTGGGGAAGAGGAGTGGTCTCCCTAATTTGTTTACTTGTAACTTCCTC 387

QY 1132 TTGTGCCCTTGGGCACTTGGCCCTTGTCTCTCAGTGTCTTCCCTTTTGACATGGGAAA 1191
    |||||
Db 386 TTGTGCCCTTGGGCACTTGGCCCTTGTCTCTCAGTGTCTTCCCTTTTGACATGGGAAA 327

QY 1192 GGAGTTGTGSCCAAAATCCCACTTCTTTCGACCTCAAGTCTGTGGCTCAGGGCTGGGG 1251
    |||||
Db 326 GGAGTTGTGSCCAAAATCCCACTTCTTTCGACCTCAAGTCTGTGGCTCAGGGCTGGGG 267

QY 1252 TGGCAGAGGAGGCGCTTACCTTATATCTGTGTGTTATCCAGGGCTCCAGACTTCTCTCC 1311
    |||||
Db 266 TGGCAGAGGAGGCGCTTACCTTATATCTGTGTGTTATCCAGGGCTCCAGACTTCTCTCC 207

QY 1312 TCTGCTGCCCCACACTGACACCTCTCTCCCTTATCTATCTCTCTGCTCCCGGCCA 1371
    |||||

```



```
/clone_lib="NIH_MGC_97"
/notes="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site 1: BamHI; Site 2: Sali-XhoI
(gtccag); Oligo-dT primed using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average
insert size 2.2 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIMH/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC Library."
```

ORIGIN

```
Query Match 35.4%; Score 532.6; DB 10; Length 751;
Best Local Similarity 96.6%; Pred. No. 4.5e-84;
Matches 598; Conservative 0; Mismatches 14; Indels 7; Gaps 5;

QY 1 CGCGATCGGACCCAGCAGTTCGGCGCGCGGAGGAGAGCGCGCGCTCAGCTCC 60
DB 107 CGCGATCGGACCCAGCAGTTCGGCGCGCGGAGGAGAGCGCGCGCTCAGCTCC 166

QY 61 TCGACCCCGCTGTCGGGCTAGTCCAGCGAGGCGGAGCGCGCTGCGGCCATGCCAGG 120
DB 167 TCGACCCCGCTGTCGGGCTAGTCCAGCGAGGCGGAGCGCGCTGCGGCCATGCCAGG 226

QY 121 CCGGCATGAGCGCGTGGCGACCGGCTGGCGCTGGTGAACGGGCGCTCGGGGGGATC 180
DB 227 CCGGCATGAGCGCGTGGCGACCGGCTGGCGCTGGTGAACGGGCGCTCGGGGGGATC 286

QY 181 GCGCGCGCGTGGCGCGCGCGCTGGTGAACGGGCGCTCGGGGGGATC 240
DB 287 GCGCGCGCGTGGCGCGCGCGCTGGTGAACGGGCGCTCGGGGGGATC 346

QY 241 ACTGTGGCAACATCGAGGAGTGGCTGTAAGTGAAGGAGTGAAGTGGTGGCGGACT 300
DB 347 ACTGTGGCAACATCGAGGAGTGGCTGTAAGTGAAGGAGTGAAGTGGTGGCGGACT 406

QY 301 TTGATCCCTACAGATGTGACCTATCAATGAAGGAGCATCTCTCCATGTTCTCAGCT 360
DB 407 TTGATCCCTACAGATGTGACCTATCAATGAAGGAGCATCTCTCCATGTTCTCAGCT 466

QY 361 ATCCGTTCTCAGCAGCGGTTGAGCATCTGCATCAACATGCTGGCTGGCGCGGCT 420
DB 467 ATCCGTTCTCAGCAGCGGTTGAGCATCTGCATCAACATGCTGGCTGGCGCGGCT 525

QY 421 GACACCTGCTCTCAGCAGCAGCAGTGG-TTGAAGGAGCATGTTCAATGTGAAGTGT 479
DB 526 GACACCTGCTCTCAGCAGCAGCAGTGG-TTGAAGGAGCATGTTCAATGTGAAGTGT 585

QY 480 GGCCTCAGCATCTGCACGCGGAGCTTACAGTCCATG-AAAGGCGGAATGTGGACG 538
DB 586 GGCCTCAGCATCTGCACGCGGAGCTTACAGTCCATGAAAGGCGGAATGTGGACG 645

QY 539 ATGGGACATCAATCAATAGATGTCTGGCCACCGAGTGTACCCCTGTCTGTGA 598
DB 646 ATGGGACATCAATCAATATATC-ATAGCATGTCTGGCCACCGAGTGTACCCCTGTGTG 701

QY 599 CCCACTTCTATAGTGCCAC 617
DB 702 ACCACTTCTATAGTGCCAC 720
```

```
RESULT 42
BI546418
LOCUS 603198719F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5260037 5',
DEFINITION mRNA sequence.
ACCESSION BI546418
VERSION BI546418.1 GI:15433730
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 714)
NIH-MGC http://img.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9apbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC Clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11655 row: f column: 06
High quality sequence stop: 704.
Location/Qualifiers
1..714
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5260037"
/tissue_type="hippocampus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_95"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site 1: BamHI; Site 2: Sali-XhoI
(gtccag); Oligo-dT primed using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average
insert size 2.5 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIMH/NHGRI, National Institutes of Health). Note: this
is a NIH_MGC Library."

FEATURES
source

ORIGIN

```
Query Match 35.2%; Score 530.2; DB 12; Length 714;  
Best Local Similarity 86.1%; Pred. No. 1.2e-83;  
Matches 678; Conservative 0; Mismatches 3; Indels 106; Gaps 4;  
QY 1 CGCGATCGGACCCAGCAGTTCGGCGCGCGGAGGAGAGCGCGCGCTCAGCTCC 60  
DB 18 CGCGATCGGACCCAGCAGTTCGGCGCGCGGAGGAGAGCGCGCGCTCAGCTCC 77  
QY 61 TCGACCCCGCTGTCGGGCTAGTCCAGCGAGGCGGAGCGCGCTGGGCCATGCCAGG 120  
DB 78 TCGACCCCGCTGTCGGGCTAGTCCAGCGAGGCGGAGCGCGCTGGGCCATGCCAGG 137  
QY 121 CCGGCATGAGCGGTTGGCGCGACCGGCTGGCTGGTGAACGGGCGCTCGGGGGGATC 180  
DB 138 CCGGCATGAGCGGTTGGCGCGACCGGCTGGCTGGTGAACGGGCGCTCGGGGGGATC 225  
QY 181 GCGCGCGCGTGGCGCGCGGCTGGTCCAGCAGGAGTGAAGGTGGTGGCTCGCCCGC 240  
DB 171 -----CTGGTCCAGCAGGAGTGAAGGTGGTGGCTCGCCCGC 209  
QY 241 ACTGTGGGCAACATCGAGGAGTGGCTGGTGAATGAAGTGCAGGCTACCCCGGACT 300  
DB 210 ACTGTGGGCAACATCG-----  
QY 301 TTGATCCCTACAGATGTGACCTATCAATGAAGGAGCATCTCTCCATGTTCTCAGCT 360  
DB 226 -----AGATGTGACCTATCAATGAAGGAGCATCTCTCCATGTTCTCAGCT 273  
QY 361 ATCCGTTCTCAGCAGCGGTTGAGCATCTGCA-TCACAAATGTGGTGGTGGCCCGGCC 419  
DB 274 ATCCGTTCTCAGCAGCAGCGGTTGAGCATCTGCA-TCACAAATGTGGTGGTGGCCCGGCC 333  
QY 420 TGACACCCCTGCTCTCAGGCGAGCAGGAGTGGTGGAGGAGCATGTTCAATGTGAACGTGT 479  
DB 334 TGACACCCCTGCTCTCAGGCGAGCAGGAGTGGTGGAGGAGCATGTTCAATGTGAACGTGT 393
```

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgabps@emall.nih.gov
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 675 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 470

FEATURES	Location/Qualifiers
source	1..528
	/organism="Homo sapiens"
	/mol_type="mrna"
	/db_xref="taxon:9606"
	/clones="IMAGE:2521953"
	/tissue_type="colonic mucosa from 3 patients with Crohn's disease"
	/lab_host="DH10B (phage-resistant)"
	/clone_lib="Soares Dieckgraefe colon NHCD"
	/notes="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCGCCGCTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Tissue samples provided by Dr. Brian Dieckgraefe (Washington University, dieck@wustl.edu), colonic mucosa represents a range of disease involvement from moderate to severe Crohn's disease; samples include both perforating (fistulas) and non-perforating samples. Library constructed by Bento Soares and M. Fatima Bonaldo."
ORIGIN	
Query Match	34.8%; Score 523.2; DB 9; Length 528;
Best Local Similarity	99.4%; Pred. No. 2.1e-82;
Matches 525; Conservative	0; Mismatches 3; Indels 0; Gaps 0;
QY	973 GATCAGGGGATACCACTTCCTGTGCCACACCCGACAGGGGCTAGAAAAATTCGTTGAGA 1032
Db	
QY	528 GATCAGGGGATACCACTTCCTGTGCCACACCCGACAGGGGCTAGAAAAATTCGTTGAGA 469
Db	
QY	1033 TTTTATATCATCTTGTCAAATTGCTTCAGTTGTAAATGTGAAAAATGGCGTCGGGAAAG 1092
Db	
QY	468 TTTTATATCATCTTGTCAAATTGCTTCAGTTGTAAATGTGAAAAATGGCGTCGGGAAAG 409
Db	
QY	1093 GAGGTGGTGTCCTAATTGTTTACTTGTGTTAACTTGTCTTGTCGCCCTGGGCACTTGGC 1152
Db	
QY	408 GAGGTGGTGTCCTAATTGTTTACTTGTGTTAACTTGTCTTGTCGCCCTGGGCACTTGGC 349
Db	
QY	1153 CTTTGTCTGCTCGAGTGCTTCCTTTGACATGGGAAAGAGTGTGGCCAAAATCCCC 1212
Db	
QY	348 CTTTGTCTGCTCGAGTGCTTCCTTTGACATGGGAAAGAGTGTGGCCAAAATCCCC 289
Db	
QY	1213 ATCTCTTGACCTCAACGTCCTGGCTCAGGCTCGGGTGGCAGAGGAGGCGCTTCACC 1272
Db	
QY	288 ATCTCTTGACCTCAACGTCCTGGCTCGGGTGGCAGAGGAGGCGCTTCACC 229
Db	
QY	1273 TTATATCTGTGTTGTTATCCAGGGCTCCAGACTTCCTCTGCTGCCCTGCCCCACCTGCACCC 1332
Db	
QY	228 TTATATCTGTGTTGTTATCCAGGGCTCCAGACTTCCTCTGCTGCCCTGCCCCACCTGCACCC 169
Db	
QY	1333 TCTCCCCCTATCTATCTCTCTCGGCTCCCGAGCCAGCTTGCTTGTCCCTC 1392
Db	
QY	168 TCTCCCCCTATCTATCTCTCTCGGCTCCCGAGCCAGCTTGCTTGTCCCTC 109
Db	
QY	1393 CTGGGGTCACTCCCTCCACTCTGACTCTGACTATGGCAGCAGAACACCAGGGCGCTGGCCCA 1452
Db	
QY	108 CTGGGGTCACTCCCTCCACTCTGACTCTGACTATGGCAGCAGAACACCAGGGCGCTGGCCCA 49
Db	
QY	1453 GTGGATTTCATGGTGATCAATTAATAAGAAAAATCGCAACCAAAAAA 1500
Db	
QY	48 GTGGATTTCATGGTGATCAATTAATAAGAAAAATCGCAACCAAAAAA 1
Db	
RESULT 45	
BE302814	
LOCUS	513 bp mRNA linear EST 14-JUL-2000
DEFINITION	ba69c04.y1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:2905638 5',
ACCESSION	mRNA sequence. BE302814

```

VERSION      BE302814.1  GI:9186562
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 513)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/;
              National Institutes of Health, Mammalian Gene Collection (MGC)
              Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: ATCC/DCTD/PTP
              cDNA Library Preparation: Ling Hong/Rubin Laboratory
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Washington University Genome Sequencing Center
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              www-bio.llnl.gov/bbrp/image/image.html
              Seq primer: -40RP from Gibco
              High quality sequence stop: 497.
              Location/Qualifiers
                1..513
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:2905638"
                  /tissue_type="melanotic melanoma"
                  /lab_host="DH10B (phage-resistant)"
                  /clone_lib="NIH MGC_20"
                  /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
                  EcoRI; cDNA made by oligo-dT priming. Directionally
                  cloned into EcoRI/XhoI sites using the following 5'
                  adaptor: GGCACGAG(G). Size-selected >500bp for average
                  insert size 1.8kb. Library constructed by Ling Hong in
                  the laboratory of Gerald M. Rubin (University of
                  California, Berkeley) using ZAP-cDNA synthesis kit
                  (Stratagene) and Superscript II RT (Life Technologies)."

```

ORIGIN	Query Match	34.0%	Score 512;	DB 10;	Length 513;
	Best Local Similarity	99.8;	Pred. No. 2e-80;		
	Matches 512;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0
QY	782	TGAAGTGTCTCAAACCCGAGGANGTGCCGAGGCTGTATTCTAGCTCTCAGCACCCCGC	841		
Db	1	TGAAGTGTCTCAAACCCGAGGATGTGCCGAGGCTGTATTCTAGCTCTCAGCACCCCGC	60		
QY	842	CACACATCCAGATTGGAGACATCCAGATGAGGCCACGGAGCAGGTGACCTTAGTGACTGT	901		
Db	61	CACACATCCAGATTGGAGACATCCAGATGAGGCCACGGAGCAGGTGACCTTAGTGACTGT	120		
QY	902	GGGAGCTCTCTTCCCTCCCAACCCCTTCATGGCTTCGCTCCCTTCGATTTTAGGT	961		
Db	121	GGGAGCTCTCTTCCCTCCCAACCCCTTCATGGCTTCGCTCCCTTCGATTTTAGGT	180		
QY	962	GTTGATTTCTGGATCAGGGATACCACTTCCTGTCCACACCCGACAGGGGCTAGAAAA	1021		
Db	181	GTTGATTTCTGGATCAGGGATACCACTTCCTGTCCACACCCGACAGGGGCTAGAAAA	240		
QY	1022	TTTGTGTTGAGATTTTATATATCATCTTGTCAAATGCTTCAGTTGTAAATGTGAAAAATGG	1081		
Db	241	TTTGTGTTGAGATTTTATATATCATCTTGTCAAATGCTTCAGTTGTAAATGTGAAAAATGG	300		
QY	1082	GCTGGGAAAGGAGGTGTCCTCAATTTGTTTTACTTTGTAACTTGTTCTGTGCCCCCT	1141		
Db	301	GCTGGGAAAGGAGGTGTCCTCAATTTGTTTTACTTTGTAACTTGTTCTGTGCCCCCT	360		
QY	1142	GGGCACTTGCGCCCTTTCTCTGCTCTCAGTGTCTCCCTTTGACATGGGAAGGAGTTCTGG	1201		
Db	361	GGGCACTTGCGCCCTTTCTCTGCTCTCAGTGTCTCCCTTTGACATGGGAAGGAGTTCTGG	420		
QY	1202	CCAAAATCCCAATCTTCCTTGGCACTCAACGTCTGTGGCTCAGGGCTGGGGTGGCAGAGG	1261		

RESULT 45	BE302814	LOCUS	BE302814	513 bp	mRNA	linear	EST 14-JUL-2000
		DEFINITION	ba69c04.y1 NIH MGC 20 Homo sapiens			CDNA clone	IMAGE:2905638 5',
			mRNA sequence.				
		ACCESSION	BE302814				


```

|||||
421 CCAAAATCCCATCTTTGACCTCAACGTCGTGGCTCAGGCTGGGTGGCGANGG 480
|||||
1262 AGGCCTTCACTTATATCTGTGTGTTATCCAG 1294
|||||
481 AGGCCTTCACTTATATCTGTGTGTTATCCAG 513
|||||

RESULT 46
BI260323      546 bp      mRNA      linear      EST 17-JUL-2001
LOCUS      602969342F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5108699 5',
DEFINITION      mRNA sequence.
ACCESSION      BI260323
VERSION      BI260323.1 GI:14818515
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 546)
AUTHORS      NIH-MGC http://mgs.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11263 row: d column: 12
High quality sequence stop: 546.
FEATURES             source
    Location/Qualifiers
        1..546
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5108699"
            /tissue_type="cervical carcinoma cell line"
            /lab_host="DH10B"
            /clone_lib="NIH_MGC_12"
            /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
            Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
            Average insert size 1.4 kb. Library prepared by Life
            Technologies."

ORIGIN
Query Match      33.9%; Score 509.8; DB 12; Length 546;
Best Local Similarity 98.3%; Pred. No. 4.9e-80;
Matches 536; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

QY      962 GTTGATTTCTGATCAGGGATACACCTTCCTGTCACACCCCGACCGGGCTAGAAAA 1021
DB      1 GTTGATTTCTGATCAGGGATACACCTTCCTGTCACACCCCGACCGGGCTAGAAAA 59
QY      1022 TTGTGTTGAGATTTTATATCATCTTGTCAAAATTGCTTCAGTTGTAATGTCAAAAATGG 1081
DB      60 TTGTGTTGAGATTTTATATCATCTTGTCAAAATTGCTTCAGTTGTAATGTCAAAAATGG 119
QY      1082 GCTGGGAAAGAGGTGGTGTCCTTAATGTTTACTTGTTAACTGTTCTTGCCCT 1141
DB      120 GCTGGGAAAGAGGTGGTGTCCTTAATGTTTACTTGTTAACTGTTCTTGCCCT 179
QY      1142 GGGCACTTGGCTTTGTCTGCTCTCAGTCTCTTCCCTTTGACATGGAAAGAGGTGTGG 1201
DB      180 GGGCACTTGGCTTTGTCTGCTCTCAGTCTCTTCCCTTTGACATGGAAAGAGGTGTGG 239
QY      1202 CCAAAATCCCATCTTTCTTGCACTCAACGTCTGTGGCTCAGGG-CTGGGGTGGCAGAGG 1260
DB      240 CCAAAATCCCATCTTTCTTGCACTCAACGTCTGTGGCTCAGGGNCTGGGGTGGCAGAGG 299

```

```

QY      1261 GAGGCTTCACTTATATCTGTGTGTTATCCAGGGCTCAGACTTCTCTCTGCCTGC 1320
DB      300 GAGGCTTCACTTATATCTGTGTGTTATCCAGGGCTCAGACTTCTCTCTGCCTGC 359
QY      1321 CCCACTGCACCCCTCTCCCTTATCTATCTCTCTTCCTGCGTCCCGACCGAGTTGGCT 1380
DB      360 CCCACTGCACCCCTCTCCCTTATCTATCTCTCTTCGGCTCCCGACCGAGTTGGCT 419
QY      1381 TCTGTCTCCCTCTCTGGGTCATCTCCCTCATCTGACTCTGACTTGGCAGCAGAACCA 1440
DB      420 TCTGTCTCCCTCTCTGGGTCATCTCCCTCATCTGACTCTGACTTGGCAGCAGAACCA 479
QY      1441 GGGCTGGCCAGTGGATTTCATGTCATCAATTAATAAAGAAAAATCGCAACCAAAAAA 1500
DB      480 GGGCTGGCCAGTGGATTTCATGTCATCAATTAATAAAGAAAAATCGCAACCAAAAAA 539
QY      1501 AAAAA 1505
DB      540 AAAAA 544

RESULT 47
BE962188/c
LOCUS      601655404R1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3846168 3',
DEFINITION      mRNA sequence.
ACCESSION      BE962188
VERSION      BE962188.2 GI:11764835
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 716)
AUTHORS      NIH-MGC http://mgs.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      On Oct 3, 2000 this sequence version replaced gi:10572893.
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLQM545 row: g column: 01
High quality sequence stop: 704.
FEATURES             source
    Location/Qualifiers
        1..716
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:3846168"
            /tissue_type="adenocarcinoma"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH_MGC_65"
            /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
            Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
            Average insert size 1.8 kb. Library constructed by Life
            Technologies."

ORIGIN
Query Match      33.7%; Score 506.6; DB 10; Length 716;
Best Local Similarity 93.2%; Pred. No. 1.8e-79;
Matches 653; Conservative 0; Mismatches 29; Indels 19; Gaps 11;

QY      823 TAGCTCTCAGACCCCGCACATCCAGATTGGAGATCCAGATGAGGCCCGGAG-881
DB      704 TAGCTCTCAGACCCCGCACATCCAGATTGG-GACATCCAGATGAGGCCCGGAG 646

```


QY	882	GCAGGTGACCTAGTGA	CTGGGAGCTCCTCTT--CCCTCCCAACCCCTTCA	TGGCTTGC	933
Db	645	GCAGGTGACCTAGTGA	CTGGGAGCTCCTCTTCCCTCGCCACCCCTT	CATGGCTTGC	586
QY	940	CTTCTGCCCTCTCGATT	TTTA-GGTGTGTATTTCTGATC	CACGGGATACACTTCTCTGT--C	996
Db	585	CTTCTGCCCTCTCGATT	TTTAGGGTGTGAAATCTGGAT	CACGGGATACACTTCTCTGTTC	526
QY	997	CACACCCCGACCA-GGGG	CTAGAAAATTTGTTTGAGATTTT	TATATCATCTTGTCAAAAT	1055
Db	525	CACACCCCGACCA	CGGGCTAGAAAATTTGTTGAGA	ATTTTATCATCTTGTCAAAAT	466
QY	1056	GCTTCAGTGTGAAA--	TGTGAAAAATGGCTGGGGAAAGAGGTGGT	GTCCCTAAATGT	1112
Db	465	GCTTCAGTGTGAAAAT	TGCAAAATTTGGCTGGGGAAAGAGGTGGT	GTCCCTAAATGT	406
QY	1113	TTTACTTGTAACTTGT	TTTCTTGTGCCCTCGGCACCTTGGCCCTT	TGTCTGTCTCAGTGTCT	1172
Db	405	TTTACTTGTAACTTGT	TTTCTTGTGCCCTCGGCACCTTGGCCCTT	TGTCTGTCTCAGTGTCT	346
QY	1173	TTCCCTTT----GACAT	GGGAAAGGAGTGTGGCCAAAATCC	CCATCTTCTTGTGCACTCA	1228
Db	345	TTCCCTATTGACAT	GGAAGGAGTGTGGCCAAAATCC	CCATCTTCTTGTGCACTCA	286
QY	1229	ACGCTGTGGCTCAGGG	CTGGGTGGCAGAGGAGGCTTCA	CCCTTAT--ATCTGTGTG	1286
Db	285	ACGCTGTGGCTCAGGG	CTGGGTGGCAGAGGAGGAGGCTT	CACCTTATAAATCTGTGTG	226
QY	1287	TTA-TCCAGGGCTCAG	ACTTCCCTCTGTGCTGCCCACTG	CACCTCTCCCGCTTATC	1345
Db	225	TTAGTCCAGGGCTCC	AGNCTTCTCTGCTTCCCTCGCC	CACTGACACCTCTCCCGCTTATC	166
QY	1346	TATCTCCTTCTGGCT	CCCCAGGCCAGTCTTGGCTTCTT	GTCCCTCTGGGCTCATCCC	1405
Db	165	TATCTCCTTCTGGCT	CCCCAGGCCAGTCTTGGCTTCTT	GTCCCTCTGGGCTCATCCC	106
QY	1406	TCCACTCTGACTCTG	ACTATGGCAGCAGAACAC	CAGGGCTTGCCCGAG-TGGATTTTCATG	1464
Db	105	TCCACTCTGACTCTG	ACTATGGCAGCAGAACAC	CAGGGCTTGCCCGAGTGGATTTTCATG	46
QY	1465	GTGATCATTA	AAAAAGAAAAATGCACCA	AAAAA	1505
Db	45	GTGATCATTA	AAAAAGAAAAATGCACCA	AAAAA	5
RESULT	48				
LOCUS	AL559035/c				
DEFINITION	AL559035 Homo sapiens T CELLS (JURKAT CELL LINE) cDNA clone CS0D010YN13 3-PRIME, mRNA sequence.	976 bp	mRNA	linear	EST 31-MAY-2003
ACCESSION	AL559035				
VERSION	AL559035.2				
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.				
TITLE	Full-length cDNA libraries and normalization				
JOURNAL	Unpublished (2001)				
COMMENT	On Feb 15, 2001 this sequence version replaced gi:12904136. Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 8845.f For more information about this cluster, see http://www.genoscope.cns.fr/ cgi-bin/cluster.cgi?seq=CS0D010CG07NP1&cluster=8845.f. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600				

```

Db      646 GAGGGGACTAGCAAGAGGCTTCGGAGGCCAGACCCATATTCGGGGCCACGTTATCTC 703
Qy      702 TCCAGGTGTGGTGGAGACACAATTCGCTTCAAATCCAGATCCAGCAAGAACCTGAGAAG 759
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      706 TCCAGGCTTGGTAGAGACACAGTTCGCTTCAAATCATGACAAACGTACCCGGGGAAG 763

RESULT 50
W58459/c  W58459          536 bp      mRNA      linear      EST 15-OCT-1996
DEFINITION
IMAGE:341723 3', mRNA sequence.

ACCESSION W58459
VERSION    W58459.1 GI:1365172
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 536)
REFERENCE  1 Hallier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares M., Tan,F.,
            Trevaaskis,E., Waterston,R., Williamson,A., Woldmann,P. and
            Wilson,R.
TITLE      The WashU-Merck EST Project
JOURNAL    Unpublished (1995)
COMMENT    Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1349 Std Error: 0.00
            Seq primer: mob.REGA+ET
            High quality sequence stop: 420.
FEATURES   Location/Qualifiers
            1..536
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="GDB:1267098"
               /db_xref="taxon:9606"
               /clone="IMAGE:341723"
               /sex="unknown"
               /dev_stage="19 weeks"
               /lab_host="DH10B (ampicillin resistant)"
               /clone_lib="Soares fetal heart NBHH19W"
               /note="organ: heart; Vector: pVT3D (Pharmacia) with a
               modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
               strand cDNA was primed with a Not I - oligo(dT) primer [5
               TGTTCACATCTGAGTGGAGCGCGCATCTTTTCTTTTCTTTT 3'],
               double-stranded cDNA was size selected, ligated to Eco RI
               adaptors (Pharmacia), digested with Not I and cloned into
               the Not I and Eco RI sites of a modified pVT3 vector
               (Pharmacia). Library went through one round of
               normalization to a Cot = 5. Library constructed by
               M.Fatima Bonaldo. This library was constructed from the
               same fetus as the fetal lung library, Soares fetal lung
               NBHH19W."

ORIGIN
Query Match      32.9%; Score 495.2; DB 14; Length 536;
Best Local Similarity 98.5%; Pred. No. 1.9e-77;
Matches 530; Conservative 0; Mismatches 5; Indels 3; Gaps 3

Qy      956 TTAGGTGTTGATTTCTGGATACCGGATACCACTTCCTGCCACACCCGACACGAGGGCT 101
Db      535 TAAAGTGTTGATTTCTGGATACCGGATACCA-TCCTGTTCACACCCGACACGAGGGCT 477
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      1016 AGAAAAATTTGTTTGAGATTTTATATCATCTTGTCAAATTCGTTCAATGTGTAATGTGAA 107

```

```

Db      476 AGAAATTTGTTGAGATTTTATATCACTCTCAAAITGCTTCAGTTGTAATGTGAA 417
QY      1076 AAATGGGCTGGGAAAGAGGTGGTGTCCCTAATCTGTTTACTTGTAACTTGTCTTGT 1135
Db      416 AAATGGGCTGGGAAAGAGGTGGTGTCCCTAATCTGTTTACTTGTAACTTGTCTTGT 357
QY      1136 GCGCTGGGCACTGTGGCCCTTGTCTCTCTCAGTGTCTTCCCTTTGACATGGGAAAGGAG 1195
Db      356 GCGCTGGGCACTGTGGCCCTTGTCTCTCTCAGTGTCTTCCCTTTGACATGGGAAAGGAG 297
QY      1196 TTGTGGCCAAATCCCATCTTCTTGCACCTCAACGTCTGTGGCTCAGGGCTGGGTGGC 1255
Db      296 TTGTGGCCAAATCCCATCTTCTTGCACCTCAACGTCTGTGGCTCAGGGCTGGGTGGC 237
QY      1256 AGAGGAGGCGCTTCACTTATATCTGTGTGTTATCCAGGGCTCCAGACTTCTCTCTG 1315
Db      236 AGAGGAGGCGCTTCACTTATATCTGTGTGTTATCCAGGGCTCCAGACTTCTCTCTG 177
QY      1316 CCTGCCCACTGCAAGCTCTCCCTTATCTATCTCTCTCGGTCCCGCCAGCTCT 1375
Db      176 CCTGCCCACTGCAAGCTCTCCCTTATCTATCTCTCTCGGTCCCGCCAGCTCT 117
QY      1376 TGGCTTCTGTCTCCCTCTGGGTGATCCCTCCACTCTGACTCTGACTATGGCAGAGAA 1435
Db      116 TGGCTTCTGT-CCCTCTGGGTGATCCCTCCACTCTGACTCTGACTATGGCAGAGAA 58
QY      1436 CACCAAGGCGCTGGCCAGTGGATTTTCATGGTGATCATTAAGAAAGAAATCGCAACC 1493
Db      57 CACCA-GGCTGGCCAGTGGATTTTCATGGTGATCATTAAGAAAGAAATCGCAACC 1

```

Search completed: September 16, 2004, 13:45:13
 Job time : 3672 secs

B/ank

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2004, 12:31:30 ; Search time 137 Seconds

(without alignments)
6096.358 Million cell updates/sec

Title: US-10-063-735-127

Perfect score: 1505

Sequence: 1 cgcggatcggaccacgag.....tcgcaaccccccaaaaaaa 1505

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 10

Total number of hits satisfying chosen parameters: 303641

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA.*

1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	445	29.6	569	4	US-09-227-357-89
2	19	1.3	608	4	US-09-370-838-172
3	19	1.3	672	4	US-09-489-039A-1571
4	19	1.3	5561	4	US-09-418-710-28
5	19	1.3	5573	4	US-09-418-710-30
6	19	1.3	5702	4	US-09-566-921-118
C 7	19	1.3	4403765	3	US-09-103-840A-2
C 8	19	1.3	4411529	3	US-09-103-840A-1
9	18	1.2	375	4	US-08-621-976-10901
10	18	1.2	1117	2	US-08-960-022-1
11	18	1.2	1278	4	US-09-489-039A-2413
12	18	1.2	1413	4	US-09-328-352-3222
C 13	18	1.2	1730	4	US-09-266-225D-7
C 14	18	1.2	2023	4	US-09-491-522-6
C 15	18	1.2	2356	4	US-09-599-360B-55
C 16	18	1.2	2450	4	US-09-491-522-2
C 17	18	1.2	2856	4	US-09-684-708A-1
C 18	18	1.2	3382	4	US-08-620-312D-449
C 19	18	1.2	3534	4	US-09-134-001C-2269
C 20	18	1.2	3915	4	US-08-023-655-1104
21	18	1.2	6669	3	US-09-212-971-5
22	18	1.2	6669	3	US-08-800-929A-5
23	18	1.2	6669	4	US-09-617-053A-5
24	18	1.2	6669	4	US-09-672-717-230
C 25	18	1.2	6692	4	US-09-491-522-1
C 26	18	1.2	11558	5	PCT-US93-06251-23
C 27	18	1.2	55298	4	US-09-491-356C-1

C 28	18	1.2	786431	4	US-09-751-389-3	Sequence 3, Appli
C 29	17	1.1	75	2	US-08-776-944-13	Sequence 13, Appli
C 30	17	1.1	146	4	US-09-621-976-16333	Sequence 16333, A
31	17	1.1	288	4	US-09-621-976-11251	Sequence 11251, A
32	17	1.1	291	4	US-09-621-976-16221	Sequence 16221, A
33	17	1.1	302	4	US-09-621-976-16866	Sequence 16866, A
C 34	17	1.1	324	4	US-09-023-655-275	Sequence 275, App
C 35	17	1.1	357	1	US-08-460-528-1	Sequence 1, Appli
C 36	17	1.1	357	2	US-08-903-224-1	Sequence 1, Appli
C 37	17	1.1	357	4	US-09-489-039A-1118	Sequence 1118, Ap
38	17	1.1	358	4	US-09-621-976-16770	Sequence 16770, A
39	17	1.1	438	4	US-09-621-976-8429	Sequence 8429, Ap
40	17	1.1	446	1	US-08-104-072B-1	Sequence 1, Appli
41	17	1.1	446	4	US-09-621-976-10841	Sequence 10841, A
C 42	17	1.1	486	4	US-09-252-991A-15808	Sequence 15808, A
C 43	17	1.1	600	4	US-09-495-050A-140	Sequence 140, App
44	17	1.1	602	4	US-09-538-709-21	Sequence 21, Appli
45	17	1.1	615	4	US-09-189-527-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1

US-09-227-357-89
; Sequence 89, Application US/09227357
; Patent No. 6342581
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: P2010P1
; CURRENT APPLICATION NUMBER: US/09/227,357
; CURRENT FILING DATE: 1999-01-08
; EARLIER APPLICATION NUMBER: PCT/US98/13684
; EARLIER FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: 60/051,926
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,793
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,925
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,929
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,803
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,732
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,931
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,932
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,916
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,930
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,918
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,920
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,733
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,795
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,919
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,928
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/055,722
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,723
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,948
; EARLIER FILING DATE: 1997-08-18

; EARLIER APPLICATION NUMBER: 60/055,949
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,953
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,950
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,947
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,964
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/056,360
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,684
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,984
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,954
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/058,785
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,664
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,660
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,661
; EARLIER FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 89
; LENGTH: 569
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-227-357-89

Query Match 29.6%; Score 445; DB 4; Length 569;
Best Local Similarity 100.0%; Pred. No. 7e-198;
Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 998 ACACCCCGACCGGGCTAGAAAATTTGTTTCAGATTTTATCATCTTGTCAATGTC 1057
Db 42 ACACCCCGACCGGGCTAGAAAATTTGTTTCAGATTTTATCATCTTGTCAATGTC 101
QY 1058 TTCAGTTGTAATGTGAATAATGGCTGGGAAAGAGGTGGTGCCTTAATGTTTAC 1117
Db 102 TTCAGTTGTAATGTGAATAATGGCTGGGAAAGAGGTGGTGCCTTAATGTTTAC 161
QY 1118 TTGTTAACTTGTCTTGCCCTGGGCACTTGGCCTTGTCTGCTCAGTGTCTCC 1177
Db 162 TTGTTAACTTGTCTTGCCCTGGGCACTTGGCCTTGTCTGCTCAGTGTCTCC 221
QY 1178 TTTCAGATGGGAAAGAGTTGGGCAAAATCCCATCTTCTTGCACCTCAACGTCGTG 1237
Db 222 TTTCAGATGGGAAAGAGTTGGGCAAAATCCCATCTTCTTGCACCTCAACGTCGTG 281
QY 1238 GCTCAGGCTGGGGTGGCAGAGGAGGCTTCACTTATATCTGTGTTGTTATCAGGCG 1297
Db 282 GCTCAGGCTGGGGTGGCAGAGGAGGCTTCACTTATATCTGTGTTGTTATCAGGCG 341
QY 1298 TCCAGATTCCTCTCTGCTGCGCCACTGCACCTCTCCCTTATCTATCTCTCTC 1357
Db 342 TCCAGATTCCTCTCTGCTGCGCCACTGCACCTCTCCCTTATCTATCTCTCTC 401
QY 1358 GCGTCCCGACCGCCAGTCTTGGCTTCTTGTCCCTCTGCGGTCATCCCTCCACTCTGACT 1417
Db 402 GCGTCCCGACCGCCAGTCTTGGCTTCTTGTCCCTCTGCGGTCATCCCTCCACTCTGACT 461
QY 1418 CTGACTATGGCAGCAACACGAG 1442
Db 462 CTGACTATGGCAGCAACACGAG 486

RESULT 2

US-09-370-838-172
; Sequence 172, Application US/09370838
; Patent No. 6444425
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Roadoh
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
; TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.475C1
; CURRENT APPLICATION NUMBER: US/09/370,838
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/285,323
; EARLIER FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 172
; LENGTH: 608
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-370-838-172

Query Match 1.3%; Score 19; DB 4; Length 608;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1131 CTTGTGCCCTGGGCACCTT 1149
Db 459 CTTGTGCCCTGGGCACCTT 477

RESULT 3

US-09-489-039A-1571
; Sequence 1571, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 1571
; LENGTH: 672
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-1571

Query Match 1.3%; Score 19; DB 4; Length 672;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 CGCAGCCGGCTGGCGCTGG 157
Db 352 CGCAGCCGGCTGGCGCTGG 370

RESULT 4

US-09-418-710-28
; Sequence 28, Application US/09418710
; Patent No. 6596482
; GENERAL INFORMATION:
; APPLICANT: Jones, Michael H.
; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATOR
; FILE REFERENCE: 06501-042001
; CURRENT APPLICATION NUMBER: US/09/418,710

```
; CURRENT FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: PCT/JP98/01783
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: JP 9/310027
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: JP 9/116570
; PRIOR FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 5561
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (346)...(4926)
US-09-418-710-28

Query Match 1.3%; Score 19; DB 4; Length 5561;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1471 ATTAAAAAGAAAAATCGC 1489
Db 4993 ATTAAAAAGAAAAATCGC 5011

RESULT 5
US-09-418-710-30
; Sequence 30, Application US/09418710
; Patent No. 6596482
; GENERAL INFORMATION:
; APPLICANT: Jones, Michael H.
; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATOR
; FILE REFERENCE: 06501-042001
; CURRENT APPLICATION NUMBER: US/09/418,710
; CURRENT FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: PCT/JP98/01783
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: JP 9/310027
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: JP 9/116570
; PRIOR FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 5573
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (346)...(4938)
US-09-418-710-30

Query Match 1.3%; Score 19; DB 4; Length 5573;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1471 ATTAAAAAGAAAAATCGC 1489
Db 5005 ATTAAAAAGAAAAATCGC 5023

RESULT 6
US-09-566-921-118
; Sequence 118, Application US/09566921
; Patent No. 6682888
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.
; APPLICANT: Tingley, Debora W.
; APPLICANT: Edwards, Carla M.
; TITLE OF INVENTION: GENES EXPRESSED IN ALZHEIMER'S DISEASE
; FILE REFERENCE: PA-0024 US
```

```
; CURRENT APPLICATION NUMBER: US/09/566,921
; CURRENT FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PERL Program
; SEQ ID NO 118
; LENGTH: 5702
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6682888 025595.15
US-09-566-921-118

Query Match 1.3%; Score 19; DB 4; Length 5702;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1131 CTTGTGCCCTGGGCACCTT 1149
Db 3087 CTTGTGCCCTGGGCACCTT 3105

RESULT 7
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match 1.3%; Score 19; DB 3; Length 4403765;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 138 GCGGACCGGCTGGCGCTG 156
Db 272023 GCGGACCGGCTGGCGCTG 272005

RESULT 8
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
```



```
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match
Best Local Similarity 1.3%; Score 19; DB 3; Length 4411529;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 138 GCGGACCGCTGGCGTG 156
Db 271911 GCGGACCGCTGGCGTG 271893

RESULT 9
US-09-621-976-10901
; Sequence 10901, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 10901
; LENGTH: 375
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-10901

Query Match
Best Local Similarity 1.2%; Score 18; DB 4; Length 375;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 952 GATTTAGGTTGATTT 969
Db 310 GATTTAGGTTGATTT 327

RESULT 10
US-08-960-022-1
; Sequence 1, Application US/08960022
; Patent No. 5976837
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: LaValle, Edward R.
; APPLICANT: Racie, Lisa A.
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; APPLICANT: Agostino, Michael J.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; TITLE OF INVENTION: ENCODING THEM
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/960,022
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Sprunger, Suzanne A.
; REGISTRATION NUMBER: 41,323
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8284
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1117 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
US-08-960-022-1

Query Match
Best Local Similarity 1.2%; Score 18; DB 2; Length 1117;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1488 GCAACCAAAAAAAAAA 1505
Db 1097 GCAACCAAAAAAAAAA 1114

RESULT 11
US-09-489-039A-2413
; Sequence 2413, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 2413
; LENGTH: 1278
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-2413

Query Match
Best Local Similarity 1.2%; Score 18; DB 4; Length 1278;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 GGTGCGGCGGCGGCGAG 37
Db 153 GGTGCGGCGGCGGCGAG 170

RESULT 12
US-09-328-352-3222
; Sequence 3222, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 3222
; LENGTH: 1413
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
```

```
US-09-328-352-3222
Query Match      1.2%; Score 18; DB 4; Length 1413;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1469 TCATTAAAAAGAAAAAT 1486
    |||||
Db 1358 TCATTAAAAAGAAAAAT 1375

RESULT 13
US-09-266-225D-7/c
; Sequence 7, Application US/09266225D
; Patent No. 6573364
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishan
; APPLICANT: Kingsmore, Stephen
; APPLICANT: Tchernev, Velizar
; TITLE OF INVENTION: Isolation and Characterization of Hermansky-Pudlak
; TITLE OF INVENTION: Syndrome (HPS) Protein Complexes and HPS Protein-
; TITLE OF INVENTION: Interacting Proteins
; FILE REFERENCE: 15966-523
; CURRENT APPLICATION NUMBER: US/09/266,225D
; CURRENT FILING DATE: 1999-03-10
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 1730
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-266-225D-7

Query Match      1.2%; Score 18; DB 4; Length 1730;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1488 GCAACCAAAAAAAAAAAAA 1505
    |||||
Db 1564 GCAACCAAAAAAAAAAAAA 1547

RESULT 14
US-09-491-522-6/c
; Sequence 6, Application US/09491522
; Patent No. 6428998
; GENERAL INFORMATION:
; APPLICANT: Colige, Alain
; APPLICANT: Lapiere, Charles M.
; APPLICANT: Prockop, Darwin J.
; TITLE OF INVENTION: RECOMBINANT N-PROTEINASE.
; TITLE OF INVENTION: AND THE PRODUCTION, METHODS AND USES THEREOF
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds, LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/491,522
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/886,333
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: Abrams, Samuel B
; REGISTRATION NUMBER: 30,605
; REFERENCE/DOCKET NUMBER: 8389-0060-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2023 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-491-522-6

Query Match      1.2%; Score 18; DB 4; Length 2023;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 23 CGCGCGCGCGCGCAGGAG 40
    |||||
Db 177 CGCGCGCGCGCGCAGGAG 160

RESULT 15
US-09-599-360B-55
; Sequence 55, Application US/09599360B
; Patent No. 6548633
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Bougueleret, L.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides
; FILE REFERENCE: GENSET.050CP3
; CURRENT APPLICATION NUMBER: US/09/599,360B
; CURRENT FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: 60/113,686
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/141,032
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 09/469,099
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: Patent.pm
; SEQ ID NO 55
; LENGTH: 2356
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 42..1097
; NAME/KEY: sig_peptide
; LOCATION: 42..110
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 4.40
; NAME/KEY: polyA_signal
; LOCATION: 2323..2328
; NAME/KEY: polyA_site
; LOCATION: 2341..2356
US-09-599-360B-55

Query Match      1.2%; Score 18; DB 4; Length 2356;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1488 GCAACCAAAAAAAAAAAAA 1505
    |||||
Db 2335 GCAACCAAAAAAAAAAAAA 2352

Search completed: September 16, 2004, 16:26:14
Job time : 149 secs
```


GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2004, 10:29:05 ; Search time 740 Seconds

(without alignments)
10265.415 Million cell updates/sec

Title: US-10-063-735-127

Perfect score: 1505

Sequence: 1 cgcggatcgaccacagcag.....tcgcaaccccaaaaaaaaa 1505

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3327077 seqs, 2523723180 residues

Total number of hits satisfying chosen parameters: 6654154

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1505	100.0	1505	13	US-10-206-915-377
2	1505	100.0	1505	13	US-10-199-670-377
3	1505	100.0	1505	13	US-10-201-858-377
4	1505	100.0	1505	13	US-10-205-890-377
5	1505	100.0	1505	13	US-10-208-024-377
6	1505	100.0	1505	13	US-10-201-853-377
7	1505	100.0	1505	13	US-10-063-745-127
8	1505	100.0	1505	13	US-10-063-512-127
9	1505	100.0	1505	13	US-10-063-513-127
10	1505	100.0	1505	13	US-10-063-569-127
11	1505	100.0	1505	13	US-10-063-551-127
12	1505	100.0	1505	13	US-10-174-581-377
13	1505	100.0	1505	13	US-10-176-483-377
14	1505	100.0	1505	13	US-10-176-749-377

15	1505	100.0	1505	13	US-10-176-914-377	Sequence 377, App
16	1505	100.0	1505	13	US-10-176-915-377	Sequence 377, App
17	1505	100.0	1505	13	US-10-063-555-127	Sequence 127, App
18	1505	100.0	1505	13	US-10-063-563-127	Sequence 127, App
19	1505	100.0	1505	13	US-10-063-594-127	Sequence 127, App
20	1505	100.0	1505	13	US-10-063-553-127	Sequence 127, App
21	1505	100.0	1505	13	US-10-063-554-127	Sequence 127, App
22	1505	100.0	1505	13	US-10-176-484-377	Sequence 377, App
23	1505	100.0	1505	13	US-10-180-550-377	Sequence 377, App
24	1505	100.0	1505	13	US-10-183-014-377	Sequence 377, App
25	1505	100.0	1505	13	US-10-187-738-377	Sequence 377, App
26	1505	100.0	1505	13	US-10-187-740-377	Sequence 377, App
27	1505	100.0	1505	13	US-10-187-883-377	Sequence 377, App
28	1505	100.0	1505	13	US-10-194-363-377	Sequence 377, App
29	1505	100.0	1505	13	US-10-194-460-377	Sequence 377, App
30	1505	100.0	1505	13	US-10-194-463-377	Sequence 377, App
31	1505	100.0	1505	13	US-10-194-484-377	Sequence 377, App
32	1505	100.0	1505	13	US-10-195-884-377	Sequence 377, App
33	1505	100.0	1505	13	US-10-195-896-377	Sequence 377, App
34	1505	100.0	1505	13	US-10-196-744-377	Sequence 377, App
35	1505	100.0	1505	13	US-10-196-755-377	Sequence 377, App
36	1505	100.0	1505	13	US-10-196-757-377	Sequence 377, App
37	1505	100.0	1505	13	US-10-197-704-377	Sequence 377, App
38	1505	100.0	1505	13	US-10-197-710-377	Sequence 377, App
39	1505	100.0	1505	13	US-10-198-758-377	Sequence 377, App
40	1505	100.0	1505	13	US-10-198-766-377	Sequence 377, App
41	1505	100.0	1505	13	US-10-199-304-377	Sequence 377, App
42	1505	100.0	1505	13	US-10-199-309-377	Sequence 377, App
43	1505	100.0	1505	13	US-10-199-313-377	Sequence 377, App
44	1505	100.0	1505	13	US-10-199-456-377	Sequence 377, App
45	1505	100.0	1505	13	US-10-201-329-377	Sequence 377, App

ALIGNMENTS

RESULT 1

US-10-206-915-377

; Sequence 377, Application US/10206915

; Publication No. US20040029221A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430R1C513

; CURRENT APPLICATION NUMBER: US/10/206,915

; CURRENT FILING DATE: 2002-07-26

; PRIOR APPLICATION NUMBER: 10/052586

; PRIOR FILING DATE: 2002-01-15

; PRIOR APPLICATION NUMBER: 60/059263

; PRIOR FILING DATE: 1997-09-18

; PRIOR APPLICATION NUMBER: 60/059266

; PRIOR FILING DATE: 1997-09-18

; PRIOR APPLICATION NUMBER: 60/062250

; PRIOR FILING DATE: 1997-10-17

; PRIOR APPLICATION NUMBER: 60/063120

; PRIOR FILING DATE: 1997-10-24

; PRIOR APPLICATION NUMBER: 60/063121

; PRIOR FILING DATE: 1997-10-24

; PRIOR APPLICATION NUMBER: 60/063486

; PRIOR FILING DATE: 1997-10-21

; PRIOR APPLICATION NUMBER: 60/063540

; PRIOR FILING DATE: 1997-10-28

```

RESULT 2
US-10-199-670-377
; Sequence 377, Application US/10199670
; Publication No. US20040033560A1
;
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
;
; TITLE OF INVENTION: SECRETED AND TRAM
;
; TITLE OF INVENTION: ACIDS ENCODING
;
; FILE REFERENCE: P3430R1C401
;
; CURRENT APPLICATION NUMBER: US/10/19
;
; CURRENT FILING DATE: 2002-07-19
;
; PRIOR APPLICATION NUMBER: 10/052586
;
; PRIOR FILING DATE: 2002-01-15
;
; PRIOR APPLICATION NUMBER: 60/059263
;
; PRIOR FILING DATE: 1997-09-18

```

```

; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION data removed - See
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 377
; LENGTH: 1505
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-199-670-377

```

Query Match	100.0%;	Score 1505;	DB 13;	Length 1505;
Best Local Similarity	100.0%;	Prod. No. 0;		
Matches 1505;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	CGCGGATCGGACCCAGACAGGTCCGGCGCGCGGCGCAGAGAGCGCCGGGGGTCTAGCTCC	60	
Db	1	CGCGGATCGGACCCAGACAGGTCCGGCGCGCGGCGCAGAGAGCGCCGGGGGTCTAGCTCC	60	
QY	61	TCGACCCCGGTGTCGGGCTAGTCAGCGAGGCGGACGGCGCGGTGGGCCCATGATGCCAAG	120	
Db	61	TCGACCCCGGTGTCGGGCTAGTCAGCGAGGCGGACGGCGCGGTGGGCCCATGATGCCAAG	120	
QY	121	CCGGCATGAGACGGTGGCGGACCGGTGCGCTGTGTGACGGGGGCGCTCGGGGGCATC	180	
Db	121	CCGGCATGAGACGGTGGCGGACCGGTGCGCTGTGTGACGGGGGCGCTCGGGGGCATC	180	
QY	181	GGCGGCGCGTGGCCCGGGCCCTGGTCCAGCAGGACTGAAGTGTGTGGGCTGGCCCGC	240	
Db	181	GGCGGCGCGTGGCCCGGGCCCTGGTCCAGCAGGACTGAAGTGTGTGGGCTGGCCCGC	240	
QY	241	ACTGTGGCAACATCGAGGAGCTGGCTGCTGAATGTAAGCTGCAGGCTACCCCGGACT	300	
Db	241	ACTGTGGCAACATCGAGGAGCTGGCTGCTGAATGTAAGCTGCAGGCTACCCCGGACT	300	
QY	301	TTGATCCCTTACAGATGTGACCTATCAATGAAGAGACATCTCTCCATGTTCTCAGCT	360	
Db	301	TTGATCCCTTACAGATGTGACCTATCAATGAAGAGACATCTCTCCATGTTCTCAGCT	360	
QY	361	ATCCGTTCTCAGCACAGCGGTGTAGACATCTGCATCAACATGTGGCTTGGCCCGGCT	420	
Db	361	ATCCGTTCTCAGCACAGCGGTGTAGACATCTGCATCAACATGTGGCTTGGCCCGGCT	420	
QY	421	GACACCTGCTCTCAGSAGCACACAGTGGTTGGAAGGACATGTTCAATGTGAACGTGCTG	480	
Db	421	GACACCTGCTCTCAGSAGCACACAGTGGTTGGAAGGACATGTTCAATGTGAACGTGCTG	480	
QY	481	GGCTCAGCATCTGCAACAGGAGGCTACAGTCCATGAAGGAGCGGAATGTGGACGAT	540	
Db	481	GGCTCAGCATCTGCAACAGGAGGCTACAGTCCATGAAGGAGCGGAATGTGGACGAT	540	
QY	541	GGGCACATCAATTAACATCAATAGCATGTCTGGGCCACCGAGTGTACCCCTGTCTGTGACC	600	
Db	541	GGGCACATCAATTAACATCAATAGCATGTCTGGGCCACCGAGTGTACCCCTGTCTGTGACC	600	
QY	601	CACCTCTATAGTCCCAACAGTATGCCGTCACTGCGCTGACAGAGGACTGAGGCAAGAG	660	
Db	601	CACCTCTATAGTCCCAACAGTATGCCGTCACTGCGCTGACAGAGGACTGAGGCAAGAG	660	
QY	661	CTTCGGAGGCCCAGACCCACATCCGAGCCACGTCATCTCTCCAGGTGTGTGGAGACA	720	

```
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C464
; CURRENT APPLICATION NUMBER: US/10/201,858
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 377
; LENGTH: 1505
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-201-858-377

Query Match      100.0%; Score 1505; DB 13; Length 1505;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  CGCGGATCGGACCCCAAGCAGGTGCGCGCGCGGCGGAGAGAGCGCGCGCGGTCTCAGCTCC 60
DB      1  CGCGGATCGGACCCCAAGCAGGTGCGCGCGCGCGGAGAGAGCGCGCGCGGTCTCAGCTCC 60

QY      61  TCGACCCCGGTTCGCGGTAGTCCAGAGGCGGACCGCGCGGTGGGCGGATGGCGAGG 120
DB      61  TCGACCCCGGTTCGCGGTAGTCCAGAGGCGGACCGCGCGGTGGGCGGATGGCGAGG 120

QY      121  CCGGCGATGGAGGCGGCGGCGGCGGCGGCTGGCGCTGGTACGCGGCGGCGGCGGCGATC 180
DB      121  CCGGCGATGGAGGCGGCGGCGGCGGCGGCGGCTGGCGCTGGTACGCGGCGGCGGCGGCGATC 180

QY      181  GCGCGGCGGCGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
DB      181  GCGCGGCGGCGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240

QY      241  ACTGTGGGCAACATCGAGAGCTGGCTGCTGAATGTGAAGTGCAGGCTACCCCGGACT 300
DB      241  ACTGTGGGCAACATCGAGAGCTGGCTGCTGAATGTGAAGTGCAGGCTACCCCGGACT 300

QY      301  TTGATCCCTTACAGATGTGACCTATCAAAATGAAGAGGACATCTCTCCATGTTCTCAGCT 360
DB      301  TTGATCCCTTACAGATGTGACCTATCAAAATGAAGAGGACATCTCTCCATGTTCTCAGCT 360

QY      361  ATCCGTTCTCAGCACAGCGGTGTAGACATCTGCATCAACAAATGCTGGTGGCGCGGCT 420
DB      361  ATCCGTTCTCAGCACAGCGGTGTAGACATCTGCATCAACAAATGCTGGTGGCGCGGCT 420

QY      421  GACACCTGCTCTCAGGAGCAGCAGTGGTGGAGGACATGTTCAATGTGAACGTGCTG 480
DB      421  GACACCTGCTCTCAGGAGCAGCAGTGGTGGAGGACATGTTCAATGTGAACGTGCTG 480

QY      481  GCCCTCAGCATCTGCACACGGGAAGCCTACCACTCCATGAAGGAGCGGAATGTGGACGAT 540
```

```
DB      481  GCCCTCAGCATCTGCACACGGGAAGCCTACCACTCCATGAAGGAGCGGAATGTGGACGAT 540
QY      541  GGGCACAATCAATTAACATAGCATGTGTGGCCACCGAGTGTACCCCTCTCTGTGACC 600
DB      541  GGGCACAATCAATTAACATAGCATGTGTGGCCACCGAGTGTACCCCTCTCTGTGACC 600
QY      601  CACTTCTATAGTGCCACCAAGTATGCCGTCTACTCGCTGACAGAGGGGACTGAGGCAAG 660
DB      601  CACTTCTATAGTGCCACCAAGTATGCCGTCTACTCGCTGACAGAGGGGACTGAGGCAAG 660
QY      661  CTTGGGAGGCGCCAGAACCCACATCCGAGCCACGTGCTCTCTCCAGGTGTGGTGAGACA 720
DB      661  CTTGGGAGGCGCCAGAACCCACATCCGAGCCACGTGCTCTCTCCAGGTGTGGTGAGACA 720
QY      721  CAATTCGGCTTCAAACTCCAGCAAGGACCTCGAGAGGCGCTGTTATCTACGTCTCAGCACCC 840
DB      721  CAATTCGGCTTCAAACTCCAGCAAGGACCTCGAGAGGCGCTGTTATCTACGTCTCAGCACCC 840
QY      841  GCACACATCCAGATTTGGAGACATCCAGATGAGGCCACCGAGCGAGGTGACCTAGTGACTG 900
DB      841  GCACACATCCAGATTTGGAGACATCCAGATGAGGCCACCGAGCGAGGTGACCTAGTGACTG 900
QY      901  TGGGAGCTCTCTCTCCCTCCACCCCTTCATGGCTTGGCTTGGCTTGGCTTGGCTTGG 960
DB      901  TGGGAGCTCTCTCTCCCTCCACCCCTTCATGGCTTGGCTTGGCTTGGCTTGGCTTGG 960
QY      961  TGTGATTTCTGGATCAGGGATACCACTCTCTGTCCACACACCCCGAGGGGCTAGAAA 1020
DB      961  TGTGATTTCTGGATCAGGGATACCACTCTCTGTCCACACACCCCGAGGGGCTAGAAA 1020
QY      1021  ATTTGTTTGGATTTTATATCATCTTGTCAAAATGCTTCAAGTGTAAATGTGAAAAATG 1080
DB      1021  ATTTGTTTGGATTTTATATCATCTTGTCAAAATGCTTCAAGTGTAAATGTGAAAAATG 1080
QY      1081  GGCTGGGAAAGAGGTGTGCTCCCTAACTGTTTACTGTTAACTGTTCTTGTGTCGCC 1140
DB      1081  GGCTGGGAAAGAGGTGTGCTCCCTAACTGTTTACTGTTAACTGTTCTTGTGTCGCC 1140
QY      1141  TGGGCACTTGGCTTTGTCTGTCTCAGTGTCTTCCCTTTGACATGGGAAAGAGTTGTG 1200
DB      1141  TGGGCACTTGGCTTTGTCTGTCTCAGTGTCTTCCCTTTGACATGGGAAAGAGTTGTG 1200
QY      1201  GCCAAAATCCCAATCTTTTGGCACTCAACGTCTGTGGCTCAGGGCTGGGGTGGCGAGG 1260
DB      1201  GCCAAAATCCCAATCTTTTGGCACTCAACGTCTGTGGCTCAGGGCTGGGGTGGCGAGG 1260
QY      1261  GAGGCTTTCACCTTATATCTGTGTTTATCCAGGGCTCCAGACTTCTCTCTGCTGCTG 1320
DB      1261  GAGGCTTTCACCTTATATCTGTGTTTATCCAGGGCTCCAGACTTCTCTCTGCTGCTG 1320
QY      1321  CCCACTGACCTCTCTCCCTTTATCTATCTCTCTCTCGGCTCCCGAGCCAGCTCTTGGCT 1380
DB      1321  CCCACTGACCTCTCTCCCTTTATCTATCTCTCTCTCGGCTCCCGAGCCAGCTCTTGGCT 1380
QY      1381  TCTTGTCCCTCTCTGGGTTCATCCCTCTCACTCTGACTGTGACTATGGGAGGACACCA 1440
DB      1381  TCTTGTCCCTCTCTGGGTTCATCCCTCTCACTCTGACTGTGACTATGGGAGGACACCA 1440
QY      1441  GGGCTTGGCCCGAGTGGATTTTCAATTAAGAAAAAAGAAAAATCGCAACCAAAAAA 1500
DB      1441  GGGCTTGGCCCGAGTGGATTTTCAATTAAGAAAAAAGAAAAATCGCAACCAAAAAA 1500
QY      1501  AAAAA 1505
DB      1501  AAAAA 1505
```



```

US-10-205-890-377
; Sequence 377, Application US/10205890
; Publication No. US200400483341
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430RIC519
; CURRENT APPLICATION NUMBER: US/10/205,890
; CURRENT FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 377
; LENGTH: 1505
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-205-890-377

Query Match      100.0%; Score 1505; DB 13; Length 1505;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  CGCGATCGGACCCAGAGTGGCGCGCGCGGAGAGAGCGCGCGGCGTCAAGTCC 60
DB      1  CGCGATCGGACCCAGAGTGGCGCGCGCGGAGAGAGCGCGCGGCGTCAAGTCC 60

QY      61  TCGACCCCGGTGTGGGCTAGTTCAGAGCGAGCGAGCGCGCGGCGTGGCCCATGCCAGG 120
DB      61  TCGACCCCGGTGTGGGCTAGTTCAGAGCGAGCGAGCGCGCGGCGTGGCCCATGCCAGG 120

QY      121  CCCGGCATGAGCGGTGGCGCGACCGGCTGGCTGGTGGCGGGGCTCGGGGGGATC 180
DB      121  CCCGGCATGAGCGGTGGCGCGACCGGCTGGCTGGTGGCGGGGCTCGGGGGGATC 180

QY      181  GCGCGCGCGGTGGCGCGGCGGCTGGTCCAGAGGACTGAAAGTGGTGGGCTGGCGCCGC 240
DB      181  GCGCGCGCGGTGGCGCGGCGGCTGGTCCAGAGGACTGAAAGTGGTGGGCTGGCGCCGC 240

QY      241  ACTGTGGCAACATCGAGGAGCTGGCTGTAATGTAAAGTGAAGTGAAGTGAAGTGAAGT 300
DB      241  ACTGTGGCAACATCGAGGAGCTGGCTGTAATGTAAAGTGAAGTGAAGTGAAGTGAAGT 300

QY      301  TTGATCCCTACAGATGTGACCTATCAAAATGAAGAGGACATCTCTCCATGTTCTCAGCT 360

```

```

DB      301  TTGATCCCTACAGATGTGACCTATCAAAATGAAGAGGACATCTCTCCATGTTCTCAGCT 360
QY      361  ATCCGTTCTCAGCACAGCGGTGTAGACATCTGCATCAACAATGCTGGCTTGGCCCGGCT 420
DB      361  ATCCGTTCTCAGCACAGCGGTGTAGACATCTGCATCAACAATGCTGGCTTGGCCCGGCT 420
QY      421  GACACCTGCTCTCAGGAGCACCGAGTGGTGGGAAGGACATGTTCAATGTGAACGTGCTG 480
DB      421  GACACCTGCTCTCAGGAGCACCGAGTGGTGGGAAGGACATGTTCAATGTGAACGTGCTG 480
QY      481  GCCCTCAGCATCTGCACACGGGAGGCTACCAAGTCCATGAAGGAGCGGAATGTGACCAT 540
DB      481  GCCCTCAGCATCTGCACACGGGAGGCTACCAAGTCCATGAAGGAGCGGAATGTGACCAT 540
QY      541  GGGCACATCAATTAACATAGCATGTCTGGCCACCGAGTGTACCCCTCTCTGTGACC 600
DB      541  GGGCACATCAATTAACATAGCATGTCTGGCCACCGAGTGTACCCCTCTCTGTGACC 600
QY      601  CACTTCTATAGTGCCACCAAGTATGCCGTCACTGGCTGACAGAGGACTGAGGCAAGAG 660
DB      601  CACTTCTATAGTGCCACCAAGTATGCCGTCACTGGCTGACAGAGGACTGAGGCAAGAG 660
QY      661  CTTGGGAGGCCAGACCCACATCCGAGCCAGTGCATCTCTCCAGGTGGTGGAGACA 720
DB      661  CTTGGGAGGCCAGACCCACATCCGAGCCAGTGCATCTCTCCAGGTGGTGGAGACA 720
QY      721  CAATTGCGCTTCAAACTCCAGCAAGGACCTCGAAGGCGAGTGCACCTATGAGCAA 780
DB      721  CAATTGCGCTTCAAACTCCAGCAAGGACCTCGAAGGCGAGTGCACCTATGAGCAA 780
QY      781  ATGAAGTGTCTCAAAACCGAGGATGTGSCCGAGGCTGTTATCTAGTCTCTCAGCACCC 840
DB      781  ATGAAGTGTCTCAAAACCGAGGATGTGSCCGAGGCTGTTATCTAGTCTCTCAGCACCC 840
QY      841  GCACACATCCAGATTTGGAGACATCCAGATGAGGCGCCAGGAGGAGTGCACCTAGTCA 900
DB      841  GCACACATCCAGATTTGGAGACATCCAGATGAGGCGCCAGGAGGAGTGCACCTAGTCA 900
QY      901  TGGAGGTCTCTCTCCCTCCCACTTTCATGGCTTGCCTTCCCTTCCCTTCCCTTCCCT 960
DB      901  TGGAGGTCTCTCTCTCCCTCCCACTTTCATGGCTTGCCTTCCCTTCCCTTCCCTTCCCT 960
QY      961  TGTGATTTCTGATATCGGATACCACTTCTCTGTCACACCCCGACAGGGGCTAGAAA 1020
DB      961  TGTGATTTCTGATATCGGATACCACTTCTCTGTCACACCCCGACAGGGGCTAGAAA 1020
QY      1021  ATTTGTTTGAGATTTTATATCATCTTGTCAAATTTGCTTCAGTTGTAATGTGAAAATG 1080
DB      1021  ATTTGTTTGAGATTTTATATCATCTTGTCAAATTTGCTTCAGTTGTAATGTGAAAATG 1080
QY      1081  GGCTGGGAAAGGAGGTGGTCCCTAAATGTTTACTTGTAACTTGTCTCTGTGCCCC 1140
DB      1081  GGCTGGGAAAGGAGGTGGTCCCTAAATGTTTACTTGTAACTTGTCTCTGTGCCCC 1140
QY      1141  TGGGCACTTGGCCCTTGTCTGCTCTAGTGTCTTCCCTTTGACATGGAAGAGGTTGTG 1200
DB      1141  TGGGCACTTGGCCCTTGTCTGCTCTAGTGTCTTCCCTTTGACATGGAAGAGGTTGTG 1200
QY      1201  GCCAAAATCCCATCTTTCGACCTCAAGTCTGTGGCTCAGGCTGGGGTGGCAGAGG 1260
DB      1201  GCCAAAATCCCATCTTTCGACCTCAAGTCTGTGGCTCAGGCTGGGGTGGCAGAGG 1260
QY      1261  GAGGCTTTCACCTTATATCTGTGTTTATCCAGGGCTCCAGACTTCTCTCTGCTGCTC 1320
DB      1261  GAGGCTTTCACCTTATATCTGTGTTTATCCAGGGCTCCAGACTTCTCTCTGCTGCTC 1320
QY      1321  CCACCTGCAACCTCTCTCCCTTATCTATCTCTCTGCTTCCGCTCCCGACCCAGTCTG 1380
DB      1321  CCACCTGCAACCTCTCTCCCTTATCTATCTCTCTGCTTCCGCTCCCGACCCAGTCTG 1380
QY      1381  TCTGTCTCCCTCTCTGGGTCATCCCTCCACTCTGACTCTGACTATGCGCAGCAGAACCA 1440

```

```
Db 1381 TCTTGTCCCTCCTGGGGTCATCCCTCCACTCTGACTCTGACTTGGCAGCAGAACCA 1440
Qy 1441 GGGCTGGCCAGTGGATTTTATGTTGATCATTTAAAAAAGAAAAATCCAAACCAAAAAA 1500
Db 1441 GGGCTGGCCAGTGGATTTTATGTTGATCATTTAAAAAAGAAAAATCCAAACCAAAAAA 1500
Qy 1501 AAAAA 1505
Db 1501 AAAAA 1505

RESULT 5
US-10-208-024-377
; Sequence 377, Application US/10208024
; Publication No. US20040048335A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C538
; CURRENT APPLICATION NUMBER: US/10/208,024
; PRIORITY FILING DATE: 2002-07-29
; PRIORITY APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIORITY APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIORITY APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIORITY APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIORITY APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIORITY APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIORITY APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIORITY APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 377
; LENGTH: 1505
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-208-024-377

Query Match 100.0%; Score 1505; DB 13; Length 1505;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGATCGGACCCAAAGCAGGTGCGCGCGCGCGGAGAGCGCGCGCGCTCAGTCC 60
Db 1 CGCGATCGGACCCAAAGCAGGTGCGCGCGCGCGGAGAGCGCGCGCGCTCAGTCC 60
Qy 61 TCGACCCCGCTCGGGCTAGTCCAGCGGCGGAGCGCGCGCTGGGCCCATGGCCAGG 120
Db 61 TCGACCCCGCTCGGGCTAGTCCAGCGGCGGAGCGCGCGCTGGGCCCATGGCCAGG 120
Qy 121 CCCGCGATGGAGCGGTGGCGCGACCGGCTGGCGCTGGTGAACGGGGGCGCTCGGGGGGCATC 180
```

```
Db 121 CCCGCGATGGAGCGGTGGCGCGACCGGCTGGCTGGTGAACGGGGGCGCTCGGGGGGCATC 180
Qy 181 GGCSCGGCCGTGGCCCGGGCCCTGGTCCAGCAGGGACTGAAGGTGGTGGGCTGGCCCCGC 240
Db 181 GGCSCGGCCGTGGCCCGGGCCCTGGTCCAGCAGGGACTGAAGGTGGTGGGCTGGCCCCGC 240
Qy 241 ACTGTGGGCAACATCGAGGAGCTGGCTGCTCAATGTAAGAGTGCAGGCTATCCCGGGACT 300
Db 241 ACTGTGGGCAACATCGAGGAGCTGGCTGCTCAATGTAAGAGTGCAGGCTATCCCGGGACT 300
Qy 301 TTGATCCCTTACAGATGTGACCTATCAAAATGAAGAGGACATCCTCTCCATGTTCTCAGCT 360
Db 301 TTGATCCCTTACAGATGTGACCTATCAAAATGAAGAGGACATCCTCTCCATGTTCTCAGCT 360
Qy 361 ATCCGTTCTCAGCAGCGGTGTAGACATCTGCGATCAACAATGCTGGCTGGCCCCGGCT 420
Db 361 ATCCGTTCTCAGCAGCGGTGTAGACATCTGCGATCAACAATGCTGGCTGGCCCCGGCT 420
Qy 421 GACACCTGCTCTCAGGCAGCACCGAGTGGTGGAGGACATGTTCAATGTGAACGTGTG 480
Db 421 GACACCTGCTCTCAGGCAGCACCGAGTGGTGGAGGACATGTTCAATGTGAACGTGTG 480
Qy 481 GCCCTCAGCATCTGCACACGGGAAGCCCTACAGTCCATGAAGAGCGGAATGTGGACGAT 540
Db 481 GCCCTCAGCATCTGCACACGGGAAGCCCTACAGTCCATGAAGAGCGGAATGTGGACGAT 540
Qy 541 GGGCAGCATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGCTGTGACC 600
Db 541 GGGCAGCATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGCTGTGACC 600
Qy 601 CACTTCTATAGTGCCACCAAGTATGCCGTCTACTGCGCTGACAGAGGGACTGAGGCAAGAG 660
Db 601 CACTTCTATAGTGCCACCAAGTATGCCGTCTACTGCGCTGACAGAGGGACTGAGGCAAGAG 660
Qy 661 CTTGGGAGGCGCCAGACCCACATCCGAGCCACGTGCATCTCTCCAGGTGTGGTGAGACA 720
Db 661 CTTGGGAGGCGCCAGACCCACATCCGAGCCACGTGCATCTCTCCAGGTGTGGTGAGACA 720
Qy 721 CAATTCGGCTTCAAACTCCAGACAGGACCCCTGAGAAAGGAGCTGCCACCTATGAGCAA 780
Db 721 CAATTCGGCTTCAAACTCCAGACAGGACCCCTGAGAAAGGAGCTGCCACCTATGAGCAA 780
Qy 781 ATGAAGTGTCTCAAAACCGAGGATGTGCCGAGGCTGTTATCTAGTCTCTCAGCACCCCC 840
Db 781 ATGAAGTGTCTCAAAACCGAGGATGTGCCGAGGCTGTTATCTAGTCTCTCAGCACCCCC 840
Qy 841 GCACATCCAGATTGGAGACATCCAGATGAGGCCCCACGGAGCAGGTGACCTAGTGACTG 900
Db 841 GCACATCCAGATTGGAGACATCCAGATGAGGCCCCACGGAGCAGGTGACCTAGTGACTG 900
Qy 901 TGGGAGCTCTCTTCCCTCCACCCTTCATGGCTTGCCCTGCTGCTTGGATTTTAGG 960
Db 901 TGGGAGCTCTCTTCCCTCCACCCTTCATGGCTTGCCCTGCTGCTTGGATTTTAGG 960
Qy 961 TGTTCATTTCTGGATCAGGGATACCACTCTCTGTCCACACCCCGACAGGGGCTAGAAA 1020
Db 961 TGTTCATTTCTGGATCAGGGATACCACTCTCTGTCCACACCCCGACAGGGGCTAGAAA 1020
Qy 1021 ATTTGTTTGGATTTTATATCATCTTGTCAAAATGCTTCAGTTGTAATGTGAAAAATG 1080
Db 1021 ATTTGTTTGGATTTTATATCATCTTGTCAAAATGCTTCAGTTGTAATGTGAAAAATG 1080
Qy 1081 GGCTGGGAAAGGAGGTGGTGTCCCTTAATTTGTTTACTGTTTAACTTCTTGTGGCCCC 1140
Db 1081 GGCTGGGAAAGGAGGTGGTGTCCCTTAATTTGTTTACTGTTTAACTTCTTGTGGCCCC 1140
Qy 1141 TGGGCACCTGGCTTTGTCTGCTCTCAGTGTCTTCCCTTTCACATGGGAAAGAGTTGTG 1200
Db 1141 TGGGCACCTGGCTTTGTCTGCTCTCAGTGTCTTCCCTTTCACATGGGAAAGAGTTGTG 1200
Qy 1201 GCCAAAAATCCCGCATCTTCTTGGACCTCAACGTCGTGTGGCTCAGGGCTGGGGTGGCAGAG 1260
Db 1201 GCCAAAAATCCCGCATCTTCTTGGACCTCAACGTCGTGTGGCTCAGGGCTGGGGTGGCAGAG 1260
```


Db 1021 ATTTGTTTGAGATTTTATATCATCTTGTCAAATTTGCTTCAGTTGTAATGTGAAAAATG 1080
Qy 1081 GGCTGGGGAAGAGGTGGTCCCTAATTTGTTTAACTTGTAACTTGTGTGCCCC 1140
Db 1081 GGCTGGGGAAGAGGTGGTCCCTAATTTGTTTAACTTGTAACTTGTGTGCCCC 1140
Qy 1141 TGGGCACTTGGCCTTTGTCTGCTCTCAGTGTCTTCCCTTGACATGGGAAAGGAGTTGTG 1200
Db 1141 TGGGCACTTGGCCTTTGTCTGCTCTCAGTGTCTTCCCTTGACATGGGAAAGGAGTTGTG 1200
Qy 1201 GCCAAATCCCACTCTTCTGACCTCAACGTCTGTGGCTCAGGGCTGGGTTGGCAGAGG 1260
Db 1201 GCCAAATCCCACTCTTCTGACCTCAACGTCTGTGGCTCAGGGCTGGGTTGGCAGAGG 1260
Qy 1261 GAGGCTTCACTTATATCTGTGTTGTTATCCAGGGCTCCAGATTCCTCTCTGCTGCTGC 1320
Db 1261 GAGGCTTCACTTATATCTGTGTTGTTATCCAGGGCTCCAGATTCCTCTCTGCTGCTGC 1320
Qy 1321 CCCAGTCCCTCTCCCTTATCTATCTCTCTGCTCCCTCCAGCCAGTCTTGGCT 1380
Db 1321 CCCAGTCCCTCTCCCTTATCTATCTCTCTGCTCCCTCCAGCCAGTCTTGGCT 1380
Qy 1381 TCTTGTCCCTCTCCCTGCTATCTCTCCACTCTGACTCTGACTATGGCAGCAGAACCA 1440
Db 1381 TCTTGTCCCTCTCCCTGCTATCTCTCCACTCTGACTCTGACTATGGCAGCAGAACCA 1440
Qy 1441 GGGCTGGCCAGTGGATTTATGTTGATCAATTAAGAAAGAAATCGCAACCAAAAAA 1500
Db 1441 GGGCTGGCCAGTGGATTTATGTTGATCAATTAAGAAAGAAATCGCAACCAAAAAA 1500
Qy 1501 AAAAA 1505
Db 1501 AAAAA 1505

RESULT 7

US-10-063-745-127
; Sequence 127, Application US/10063745
; Publication No. US20040058411A1
; GENERAL INFORMATION:
; APPLICANT: Baton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT FILING DATE: 2002-05-09
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 127
; LENGTH: 1505
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-063-745-127

Query Match 100.0%; Score 1505; DB 13; Length 1505;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGATCGGACCAAGCAGGTGGGGCGGCGGCGGAGGAGCGGCGGCGGTCAAGTCC 60
Db 1 CGGGATCGGACCAAGCAGGTGGGGCGGCGGCGGAGGAGCGGCGGCGGTCAAGTCC 60
Qy 61 TCGACCCCGTGTGGGCTAGTCCAGCGAGCGGAGCGGGCGGCGGTGGCCAGG 120
Db 61 TCGACCCCGTGTGGGCTAGTCCAGCGAGCGGCGGCGGCGGTGGCCAGG 120

Qy 121 CCCGSCATGGAGCGTGGCGGACCGGCTGGCGCTGGTGACGGGGGCTTCGCGGGGCATC 180
Db 121 CCCGSCATGGAGCGTGGCGGACCGGCTGGCGCTGGTGACGGGGGCTTCGCGGGGCATC 180
Qy 181 GGGCGGGCGGTGGCGGCGGCGGCTGGTCCAGCAGGAGCTGAAGGTGGTGGGCTGGCGCGC 240
Db 181 GGGCGGGCGGTGGCGGCGGCGGCTGGTCCAGCAGGAGCTGAAGGTGGTGGGCTGGCGCGC 240
Qy 241 ACTGTGGGCAACATCGAGGAGCTGGCTGCTCAATGTAAAGTGCAGGCTACCCGGGACT 300
Db 241 ACTGTGGGCAACATCGAGGAGCTGGCTGCTCAATGTAAAGTGCAGGCTACCCGGGACT 300
Qy 301 TTGATCCCTACAGATGTACCTATCAAAATGAAGAGGACATCCCTCTCCATGTTCTCAGCT 360
Db 301 TTGATCCCTACAGATGTACCTATCAAAATGAAGAGGACATCCCTCTCCATGTTCTCAGCT 360
Qy 361 ATCCGTTCTCAGCAGCGGTGTAGACATCTGCGATCAACAATGCTGGCTGGCCGGGCT 420
Db 361 ATCCGTTCTCAGCAGCGGTGTAGACATCTGCGATCAACAATGCTGGCTGGCCGGGCT 420
Qy 421 GACACCTGCTCTCAGGCGAGCAGTGGTGGAGGACATGTTCAATGTCAACGTGTG 480
Db 421 GACACCTGCTCTCAGGCGAGCAGTGGTGGAGGACATGTTCAATGTCAACGTGTG 480
Qy 481 GCCCTCAGCATCTGCACACGGGAAAGCCTACCAAGTCCATGAAGGAGCGGAATGTGGACAT 540
Db 481 GCCCTCAGCATCTGCACACGGGAAAGCCTACCAAGTCCATGAAGGAGCGGAATGTGGACAT 540
Qy 541 GGGCACTCATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGTCTGACC 600
Db 541 GGGCACTCATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGTCTGACC 600
Qy 601 CACTTCTATAGTGCACCAAGTATGCGCTCACTGCGCTGACAGAGGAGCTGAGGCAAGAG 660
Db 601 CACTTCTATAGTGCACCAAGTATGCGCTCACTGCGCTGACAGAGGAGCTGAGGCAAGAG 660
Qy 661 CTTGGGAGGCGCCAGACCCACATCCGAGCCAGTGCATCTCTCCAGGTGTGGTGAGACA 720
Db 661 CTTGGGAGGCGCCAGACCCACATCCGAGCCAGTGCATCTCTCCAGGTGTGGTGAGACA 720
Qy 721 CAATTGCGCTTCAAACTCCACGACCAAGACCTCGAGAGGAGCTGCCACCTATGAGCAA 780
Db 721 CAATTGCGCTTCAAACTCCACGACCAAGACCTCGAGAGGAGCTGCCACCTATGAGCAA 780
Qy 781 ATGAAGTGTCTCAAAACCCGAGGATGTGGCCGAGGCTGTTATCTAAGTGTGGTGGAGCA 840
Db 781 ATGAAGTGTCTCAAAACCCGAGGATGTGGCCGAGGCTGTTATCTAAGTGTGGTGGAGCA 840
Qy 841 GCACACATCCAGATTGGAGACATCCAGATGAGGCGCCACGAGCAGGTGACCTAGTACATG 900
Db 841 GCACACATCCAGATTGGAGACATCCAGATGAGGCGCCACGAGCAGGTGACCTAGTACATG 900
Qy 901 TGGGAGCTCCTCTTCCCTCCACCTTCATGGCTGGCTCCTGCTCTGATTTAGG 960
Db 901 TGGGAGCTCCTCTTCCCTCCACCTTCATGGCTGGCTCCTGCTCTGATTTAGG 960
Qy 961 TGTGATTTCTGATCAGCGGATACCACTTCTGTCACACCCCGAGGAGCTAGAAA 1020
Db 961 TGTGATTTCTGATCAGCGGATACCACTTCTGTCACACCCCGAGGAGCTAGAAA 1020
Qy 1021 ATTTGTTTGAGATTTTATATCATCTTGTCAAATTTGCTTCAGTTGTAATGTGAAAAATG 1080
Db 1021 ATTTGTTTGAGATTTTATATCATCTTGTCAAATTTGCTTCAGTTGTAATGTGAAAAATG 1080
Qy 1081 GGTGGGGAAGAGGTGGTGGCTTAAATGTTTACTTGTAACTGTTCTGTGCCCC 1140
Db 1081 GGTGGGGAAGAGGTGGTGGCTTAAATGTTTACTTGTAACTGTTCTGTGCCCC 1140
Qy 1141 TGGGCACTTGGCCTTTGTCTGCTCTCAGTGTCTTCCCTTTGACATGGGAAAGGAGTTGTG 1200
Db 1141 TGGGCACTTGGCCTTTGTCTGCTCTCAGTGTCTTCCCTTTGACATGGGAAAGGAGTTGTG 1200

```
QY 1201 GCCAAATCCCATCTTCTTGACCTCAACGCTGTGGCTCAGGCTGGGTGGCAGAG 1260
Db 1201 GCCAAATCCCATCTTCTTGACCTCAACGCTGTGGCTCAGGCTGGGTGGCAGAG 1260
QY 1261 GAGGCTTCACTTATATCTGTGTGTATCCAGGCTCCAGACTTCTCTGCTGC 1320
Db 1261 GAGGCTTCACTTATATCTGTGTGTATCCAGGCTCCAGACTTCTCTGCTGC 1320
QY 1321 CCCACTGACCTCTCCCTTATCTATCTCTCTGCTCCAGCTCCAGCTTGGCT 1380
Db 1321 CCCACTGACCTCTCCCTTATCTATCTCTCTGCTCCAGCTCCAGCTTGGCT 1380
QY 1381 TCTTGTCCCTCTGCGGTTCATCCCTCCACTCTGACTCTGACTATGGCAGAGAACCA 1440
Db 1381 TCTTGTCCCTCTGCGGTTCATCCCTCCACTCTGACTCTGACTATGGCAGAGAACCA 1440
QY 1441 GGGCTGGCCCACTGATTTTCAATGATGATCAATTAATAAAGAAAAATCGCAACAAAA 1500
Db 1441 GGGCTGGCCCACTGATTTTCAATGATGATCAATTAATAAAGAAAAATCGCAACAAAA 1500
QY 1501 AAAAA 1505
Db 1501 AAAAA 1505

RESULT 8
US-10-063-512-127
; Sequence 127, Application US/10063512
; Publication No. US20030018183A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,512
; CURRENT FILING DATE: 2002-05-01
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 127
; LENGTH: 1505
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-063-512-127

Query Match 100.0%; Score 1505; DB 13; Length 1505;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGATCGGACCCAGCAGTGTGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTCAGCTCC 60
Db 1 CGCGATCGGACCCAGCAGTGTGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTCAGCTCC 60
QY 61 TCGACCCCGGTGTGGGTGTGTTCAGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
Db 61 TCGACCCCGGTGTGGGTGTGTTCAGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
QY 121 CCGCGCATGGAGCGGTGTGGGTGTGTTCAGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
Db 121 CCGCGCATGGAGCGGTGTGGGTGTGTTCAGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
QY 181 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
Db 181 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
QY 241 ACTGTGGCAACATCGAGGAGCTGGCTGCTGAATGTAAGAGTGCAGGCTACCCCGGACT 300
```

```
Db 241 ACTGTGGCAACATCGAGGAGCTGGCTGCTGAATGTAAGAGTGCAGGCTACCCCGGACT 300
QY 301 TTGATCCCCCTCAGATGTGACCTATCAAAATGAAGAGGACATCTCTCCATGTTCTCAGCT 360
Db 301 TTGATCCCCCTCAGATGTGACCTATCAAAATGAAGAGGACATCTCTCCATGTTCTCAGCT 360
QY 361 ATCCCTTCTCAGCA CAGCGGTGTAGCATCTGCATCAACAATGCTGGTGTGCGCGCT 420
Db 361 ATCCCTTCTCAGCA CAGCGGTGTAGCATCTGCATCAACAATGCTGGTGTGCGCGCT 420
QY 421 GACACCTCTCTCAGGCGGCGGCTGAGCATCTGCATCAACAATGCTGGTGTGCGCGCT 480
Db 421 GACACCTCTCTCAGGCGGCGGCTGAGCATCTGCATCAACAATGCTGGTGTGCGCGCT 480
QY 481 GCCCTCAGCATCTGCA CAGCGGAGCTACCACTGAGGAGCGGAGTGTGACCAT 540
Db 481 GCCCTCAGCATCTGCA CAGCGGAGCTACCACTGAGGAGCGGAGTGTGACCAT 540
QY 541 GGGCAGCATTTAATCAATAGCATGTGTGCGCA CCGAGTGTACCCCTGTCTGTGACC 600
Db 541 GGGCAGCATTTAATCAATAGCATGTGTGCGCA CCGAGTGTGTACCCCTGTCTGTGACC 600
QY 601 CACTTCTATAGTGGCCACCAAGTATGCGCTCACTGCGCTGACAGGGAGTGGGCAAGAG 660
Db 601 CACTTCTATAGTGGCCACCAAGTATGCGCTCACTGCGCTGACAGGGAGTGGGCAAGAG 660
QY 661 CTTGGGAGGCGGCGGAGCCACATCCGAGCCAGCTGATCTCTCCAGGTGTGGTGAGACA 720
Db 661 CTTGGGAGGCGGCGGAGCCACATCCGAGCCAGCTGATCTCTCCAGGTGTGGTGAGACA 720
QY 721 CAATTCGCTTCAAACTCCAGCAAGGAGCCCTGAGAGGCGAGTGCACCTATGAGCAA 780
Db 721 CAATTCGCTTCAAACTCCAGCAAGGAGCCCTGAGAGGCGAGTGCACCTATGAGCAA 780
QY 781 ATGAAGTGTCTCAAACTCCAGGAGTGTGGCGGAGGCTGTATCTAGCTCTCAGCACCCCT 840
Db 781 ATGAAGTGTCTCAAACTCCAGGAGTGTGGCGGAGGCTGTATCTAGCTCTCAGCACCCCT 840
QY 841 GCACATCCAGATTTGAGAGATCCAGATGAGGCGGCGGAGGAGTGCACCTAGTGAAGT 900
Db 841 GCACATCCAGATTTGAGAGATCCAGATGAGGCGGCGGAGGAGTGCACCTAGTGAAGT 900
QY 901 TGGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 960
Db 901 TGGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 960
QY 961 TGTGATTTCTGGATCAGGATACCACTTCTCTCCACACCCGAGGAGTGCACCTAGTGAAGT 1020
Db 961 TGTGATTTCTGGATCAGGATACCACTTCTCTCTCCACACCCGAGGAGTGCACCTAGTGAAGT 1020
QY 1021 ATTTGTTTGTAGATTTTATATATCTTGTCAAAATGCTTCAAGTGTGAATGTGAATAATG 1080
Db 1021 ATTTGTTTGTAGATTTTATATATCTTGTCAAAATGCTTCAAGTGTGAATGTGAATAATG 1080
QY 1081 GGTGCGGAAAGAGGAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1140
Db 1081 GGTGCGGAAAGAGGAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1140
QY 1141 TGGGCACTTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1200
Db 1141 TGGGCACTTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1200
QY 1201 GCCAAAAATCCCATCTTCTTGTGACCTCAACGCTGTGGCTCAGGCTGGGTGGCAGAGG 1260
Db 1201 GCCAAAAATCCCATCTTCTTGTGACCTCAACGCTGTGGCTCAGGCTGGGTGGCAGAGG 1260
QY 1261 GAGGCTTCACTTATATCTGTGTGTATCCAGGCTCCAGACTTCTCTCTGCTGC 1320
Db 1261 GAGGCTTCACTTATATCTGTGTGTATCCAGGCTCCAGACTTCTCTCTGCTGC 1320
QY 1321 CCCACTGACCTCTCCCTTATCTATCTCTCTGCTCCAGCTCCAGCTTGGCT 1380
```

```
Db 1321 CCACCTGCACCCCTCTCCCTTATCTATCTCTCTCTCGGCTCCCGAGCCAGCTCTTGGCT 1380
Qy 1381 TCTTGTCCCTCTCTGGGTCTATCTCTCACTCTGACTCTGACTATGGCAGCAGAACACA 1440
Db 1381 TCTTGTCCCTCTCTGGGTCTATCTCTCACTCTGACTCTGACTATGGCAGCAGAACACA 1440
Qy 1441 GGGCTGCGCCAGTGGATTTTCATGCTGATCATTAAGAAAGAAAAATCGAACCAAAAAA 1500
Db 1441 GGGCTGCGCCAGTGGATTTTCATGCTGATCATTAAGAAAGAAAAATCGAACCAAAAAA 1500
Qy 1501 AAAAA 1505
Db 1501 AAAAA 1505

RESULT 9
US-10-063-513-127
; Sequence 127, Application US/10063513
; Publication No. US20030018172A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,513
; PRIORITY FILING DATE: 2002-05-01
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 127
; LENGTH: 1505
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-063-513-127

Query Match 100.0%; Score 1505; DB 13; Length 1505;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGGATCGGACCCCAAGCAGGTCTGGCGCGCGCGCAGGAGCGCGCGGCGCTCAGCTCC 60
Db 1 CGCGGATCGGACCCCAAGCAGGTCTGGCGCGCGCGCAGGAGCGCGCGGCGCTCAGCTCC 60
Qy 61 TCGACCCCGGTCTGGGCTAGTCTCAGCGAGGCGGAGCGGCGCGCTGGGCGCATGGCCAGG 120
Db 61 TCGACCCCGGTCTGGGCTAGTCTCAGCGAGGCGGAGCGGCGCGCTGGGCGCATGGCCAGG 120
Qy 121 CCGGCGATGAGCGGTGGCGCAGCGGTGGCGCTGGTGACGGGGCGCTCGGGGGGCATC 180
Db 121 CCGGCGATGAGCGGTGGCGCAGCGGTGGCGCTGGTGACGGGGCGCTCGGGGGGCATC 180
Qy 181 GCGCGGCGGTGGCGCGGCGCTGGTCCAGCGGAGCTGAAGGTGGTGGGCTGGCGCGCG 240
Db 181 GCGCGGCGGTGGCGCGGCGCTGGTCCAGCGGAGCTGAAGGTGGTGGGCTGGCGCGCG 240
Qy 241 ACTGTGGGCAACATCGAGAGCTGGTCTGATGTAGAGTGCAGGCTACCCCGGACT 300
Db 241 ACTGTGGGCAACATCGAGAGCTGGTCTGATGTAGAGTGCAGGCTACCCCGGACT 300
Qy 301 TTGATCCCCCTACAGATGTGACCTTATCAAAATGAAGAGGACATCTCTCCAATGTTCTCAGCT 360
Db 301 TTGATCCCCCTACAGATGTGACCTTATCAAAATGAAGAGGACATCTCTCCAATGTTCTCAGCT 360
Qy 361 ATCCGTTCTCAGCACAGCGGTGTAGACATCTGCATCAACAATGCTGGCTTGGCCCGGCT 420
Db 361 ATCCGTTCTCAGCACAGCGGTGTAGACATCTGCATCAACAATGCTGGCTTGGCCCGGCT 420
```

```
Qy 421 GACACCTGTCTCTCAGCGACACACAGTGGTTGGAGGACATGTTCAATGTGAACGTGCTG 480
Db 421 GACACCTGTCTCTCAGCGACACACAGTGGTTGGAGGACATGTTCAATGTGAACGTGCTG 480
Qy 481 GCCCTCAGCATCTGCACACGGGAAGCCTTACAGTCCATGAAGAGCGGAATGTGGACGAT 540
Db 481 GCCCTCAGCATCTGCACACGGGAAGCCTTACAGTCCATGAAGAGCGGAATGTGGACGAT 540
Qy 541 GGGCACATCATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGTCTGTGACC 600
Db 541 GGGCACATCATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGTCTGTGACC 600
Qy 601 CACTTCTATAGTGGCACCAAGTATGCCGTCTACTGCGTGCAGAGGAGCTCAGGCAAGAG 660
Db 601 CACTTCTATAGTGGCACCAAGTATGCCGTCTACTGCGTGCAGAGGAGCTCAGGCAAGAG 660
Qy 661 CTTGGGAGGCGCCAGACCCACATCCGAGCCAGTGCATCTCTCCAGGTGTGGTGGAGACA 720
Db 661 CTTGGGAGGCGCCAGACCCACATCCGAGCCAGTGCATCTCTCCAGGTGTGGTGGAGACA 720
Qy 721 CAATTCGCTTCAAACTCCAGACCAAGACCCCTGAGAAGGAGCTGCCACCTATGAGCAA 780
Db 721 CAATTCGCTTCAAACTCCAGACCAAGACCCCTGAGAAGGAGCTGCCACCTATGAGCAA 780
Qy 781 ATGAAGTGTCTCAAAACCGAGGATGTGGCGGAGGTGTATCTACGTCTCTCAGCACCCCC 840
Db 781 ATGAAGTGTCTCAAAACCGAGGATGTGGCGGAGGTGTATCTACGTCTCTCAGCACCCCC 840
Qy 841 GCACACATCCAGATTTGGAGACATCCAGATGAGGCCACCGAGCAGGTGACCTTAGTGACTG 900
Db 841 GCACACATCCAGATTTGGAGACATCCAGATGAGGCCACCGAGCAGGTGACCTTAGTGACTG 900
Qy 901 TGGGAGCTCTCTCTCCCTCCCGACCTTCAATGCTTGGCTTGGCTTGGCTTGGCTTGGCT 960
Db 901 TGGGAGCTCTCTCTCTCCCTCCCGACCTTCAATGCTTGGCTTGGCTTGGCTTGGCTTGGCT 960
Qy 961 TGTGATTTCTGGATCAGGGATACCACTTCTGTCTCCACACCCCGACCGAGGGCTAGAAA 1020
Db 961 TGTGATTTCTGGATCAGGGATACCACTTCTGTCTCCACACCCCGACCGAGGGCTAGAAA 1020
Qy 1021 ATTTGTTTGGATTTTATATCATCTTGTCAAAATGCTTCAATGTTGTTGTTGTTGTTG 1080
Db 1021 ATTTGTTTGGATTTTATATCATCTTGTCAAAATGCTTCAATGTTGTTGTTGTTGTTG 1080
Qy 1081 GGTCTGGGAAAGAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1140
Db 1081 GGTCTGGGAAAGAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1140
Qy 1141 TGGGCATTTGGCTTTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 1200
Db 1141 TGGGCATTTGGCTTTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 1200
Qy 1201 GCGAAAATCCCATCTTCTTGCACCTCAACGTCTGTGGCTCAGGCTGGGGTGGCAGAGG 1260
Db 1201 GCGAAAATCCCATCTTCTTGCACCTCAACGTCTGTGGCTCAGGCTGGGGTGGCAGAGG 1260
Qy 1261 GAGGCTTTCACCTTATATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1320
Db 1261 GAGGCTTTCACCTTATATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1320
Qy 1321 CCCACTGCAACCTCTCCCTCTTATCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1380
Db 1321 CCCACTGCAACCTCTCCCTCTTATCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1380
Qy 1381 TCTGTGCCCTCTCTGGGTCTATCCCTCCACTCTGACTCTGACTATGGCAGCAGAACACA 1440
Db 1381 TCTGTGCCCTCTCTGGGTCTATCCCTCCACTCTGACTCTGACTATGGCAGCAGAACACA 1440
Qy 1441 GGGCGCTGGCCAGTGGATTTTCATGCTGATCAATTAAGAAAGAAAAATCGCAACCAAAAAA 1500
Db 1441 GGGCGCTGGCCAGTGGATTTTCATGCTGATCAATTAAGAAAGAAAAATCGCAACCAAAAAA 1500
```


QY 1501 AAAAA 1505
 Db 1501 AAAAA 1505

RESULT 10
 US-10-063-569-127
 ; Sequence 127, Application US/10063569
 ; Publication No. US20030018168A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Eaton, Dan L.
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, Christopher J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; TITLE OF INVENTION: ACIDS ENCODING THE SAME
 ; FILE REFERENCE: P3230R1C1
 ; CURRENT APPLICATION NUMBER: US/10/063,569
 ; CURRENT FILING DATE: 2002-05-02
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 170
 ; SEQ ID NO 127
 ; LENGTH: 1505
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 US-10-063-569-127

Query Match 100.0%; Score 1505; DB 13; Length 1505;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGGATCGGACCAAGCAGGTGCGCGCGCGCGAGAGCGCGGGCGTCAAGTCC 60
 Db 1 CGCGGATCGGACCAAGCAGGTGCGCGCGCGCGAGAGCGCGGGCGTCAAGTCC 60

QY 61 TCGACCCCGGTGTCGGGCTAGTCAGCGAGCGGCGCGCGCGCGCGCGCGCGCG 120
 Db 61 TCGACCCCGGTGTCGGGCTAGTCAGCGAGCGGCGCGCGCGCGCGCGCGCG 120

QY 121 CCCGGCATGAGCGGTGCGCGAGCGGCTGCGCGTGTGAGCGGGCGCTCGGGGGGATC 180
 Db 121 CCCGGCATGAGCGGTGCGCGAGCGGCTGCGCGTGTGAGCGGGCGCTCGGGGGGATC 180

QY 181 GCGCGCGCGGTGCG 240
 Db 181 GCGCGCGCGGTGCG 240

QY 241 ACTGTGGGCAACATCGAGGAGTGGCTGCTGCTGAATGTAAGAGTGAGGCTACCCCGGACT 300
 Db 241 ACTGTGGGCAACATCGAGGAGTGGCTGCTGCTGAATGTAAGAGTGAGGCTACCCCGGACT 300

QY 301 TTGATCCCTCAGATGACATCAATGAAGAGGACATCCCTCTCCATGTTCTCAGCT 360
 Db 301 TTGATCCCTCAGATGACATCAATGAAGAGGACATCCCTCTCCATGTTCTCAGCT 360

QY 361 ATCCGTTCTCAGCAGCGGTGAGACATCTGACATCAACATGCTGGCTTGGCCCGGCT 420
 Db 361 ATCCGTTCTCAGCAGCGGTGAGACATCTGACATCAACATGCTGGCTTGGCCCGGCT 420

QY 421 GACACCTGCTCTCAGGCGAGCAGGCTGGTGGAGGACATGTTCAATGTGAAGTGGCTG 480
 Db 421 GACACCTGCTCTCAGGCGAGCAGGCTGGTGGAGGACATGTTCAATGTGAAGTGGCTG 480

QY 481 GCCCTCAGCATCTGCACACGGGAGCCTACAGTCCATGAAGGAGCGGAATGTGGACAT 540
 Db 481 GCCCTCAGCATCTGCACACGGGAGCCTACAGTCCATGAAGGAGCGGAATGTGGACAT 540

QY 541 GGGCACATCAATCAATAGCATGTCTGCGCCAGCGAGTGTACCCCTGTCTGTGACC 600

Db 541 GGGCACATCAATCAATAGCATGTCTGCGCCACCGAGTGTACCCCTGTCTGTGACC 600
 QY 601 CACTTCTATAGTGCACCAAGTATGCCGTCACTCGCTGACAGAGGAGCTGAGGCAAGAG 660
 Db 601 CACTTCTATAGTGCACCAAGTATGCCGTCACTCGCTGACAGAGGAGCTGAGGCAAGAG 660
 QY 661 CTTGCGGAGGCCAGACCCACATCCGAGCCAGTGCACTCTCCAGGTGTGGTGGAGACA 720
 Db 661 CTTGCGGAGGCCAGACCCACATCCGAGCCAGTGCACTCTCCAGGTGTGGTGGAGACA 720
 QY 721 CAATTCCGCTTCAAACCTCCAGCAAGAGCCCTCAGAGGCGAGCTGCCACCTATGAGCAA 780
 Db 721 CAATTCCGCTTCAAACCTCCAGCAAGAGCCCTCAGAGGCGAGCTGCCACCTATGAGCAA 780
 QY 781 ATGAAGTGTCTCAAACCCGAGGATGTGCGCGAGGCTGTATCTACGTCTCAGCACCCCC 840
 Db 781 ATGAAGTGTCTCAAACCCGAGGATGTGCGCGAGGCTGTATCTACGTCTCAGCACCCCC 840
 QY 841 GCACACATCCAGATTGGAGACATCCAGATGAGGCGCCACGGAGCGAGTGAACCTAGTGA 900
 Db 841 GCACACATCCAGATTGGAGACATCCAGATGAGGCGCCACGGAGCGAGTGAACCTAGTGA 900
 QY 901 TGGGAGCTCTCTCCCTCCCTCCCACTTCATGCGCTTCCCTCCCTCTGGATTTTAGG 960
 Db 901 TGGGAGCTCTCTCCCTCCCTCCCACTTCATGCGCTTCCCTCCCTCTGGATTTTAGG 960
 QY 961 TGTGATTTCTGGATCACGGGATACCACTTCTCTCCACACCCCGAGGGCTAGAAA 1020
 Db 961 TGTGATTTCTGGATCACGGGATACCACTTCTCTCCACACCCCGAGGGCTAGAAA 1020
 QY 1021 ATTTGTTTGAGATTTTATATCATCTTGTGTAATGCTTCAGTTGTAAATGTGAAAAATG 1080
 Db 1021 ATTTGTTTGAGATTTTATATCATCTTGTGTAATGCTTCAGTTGTAAATGTGAAAAATG 1080
 QY 1081 GGCTGGGGAAGAGGAGTGGTCCCTAAATGTTTACTTGTAACTTGTCTTGTGCCCC 1140
 Db 1081 GGCTGGGGAAGAGGAGTGGTCCCTAAATGTTTACTTGTAACTTGTCTTGTGCCCC 1140
 QY 1141 TGGGCACCTTGGCCCTTGTCTGCTCTCAGTGTCTTCCCTTTGACATGGGAAAGAGTTGTG 1200
 Db 1141 TGGGCACCTTGGCCCTTGTCTGCTCTCAGTGTCTTCCCTTTGACATGGGAAAGAGTTGTG 1200
 QY 1201 GCAAAATCCCATCTTCTTGGACCTCAAGTCTGTGGCTCAGGGCTGGGGTGGCAGAGG 1260
 Db 1201 GCAAAATCCCATCTTCTTGGACCTCAAGTCTGTGGCTCAGGGCTGGGGTGGCAGAGG 1260
 QY 1261 GAGGCTTCCACCTTATATCTGTGTTTATCCAGGGCTCCAGACTTCTCTCTGCGCTGC 1320
 Db 1261 GAGGCTTCCACCTTATATCTGTGTTTATCCAGGGCTCCAGACTTCTCTCTGCGCTGC 1320
 QY 1321 CCCACTGCACCCCTCTCCCCCTTATCTATCTCTCGGCTCCCGAGCCAGCTTGGCT 1380
 Db 1321 CCCACTGCACCCCTCTCCCCCTTATCTATCTCTCGGCTCCCGAGCCAGCTTGGCT 1380
 QY 1381 TCTTGTCCCTCTCGGGGTATCTCTCCACTCTGACTCTGACTATGGCAGAGAACCA 1440
 Db 1381 TCTTGTCCCTCTCGGGGTATCTCTCCACTCTGACTCTGACTATGGCAGAGAACCA 1440
 QY 1441 GGGGCTGGCCCGAGTGGATTTTCTGGTGTATTAATAAAGAAAAATCGCAACCAAAAAA 1500
 Db 1441 GGGGCTGGCCCGAGTGGATTTTCTGGTGTATTAATAAAGAAAAATCGCAACCAAAAAA 1500
 QY 1501 AAAAA 1505
 Db 1501 AAAAA 1505

RESULT 11
 US-10-063-551-127
 ; Sequence 127, Application US/10063551
 ; Publication No. US2002018349A1
 ; GENERAL INFORMATION:

; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,551
; CURRENT FILING DATE: 2002-05-02
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 127
; LENGTH: 1505
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-063-551-127

Query Match 100.0%; Score 1505; DB 13; Length 1505;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	CGCGGATCGGACCAAGCAGGTGCGCGCGCGCGCAGGAGAGCGCGCGGGCGTCAGCTCC	60
Db	1	CGCGGATCGGACCAAGCAGGTGCGCGCGCGCGCAGGAGAGCGCGCGGGCGTCAGCTCC	60
Qy	61	TCGACCCCGGTGTCGGGTAGTCCAGAGGAGCGGCGCGGTGGGCCCATGGCCAGG	120
Db	61	TCGACCCCGGTGTCGGGTAGTCCAGAGGAGCGGCGCGGTGGGCCCATGGCCAGG	120
Qy	121	CCCGGATGAGAGCGGTGCGCGAGCGCGTGGCTGCGTGGTGGCGGCGCTCGGGGGGCATC	180
Db	121	CCCGGATGAGAGCGGTGCGCGAGCGCGTGGCTGCGTGGTGGCGGCGCTCGGGGGGCATC	180
Qy	181	GGCGGCGGTGGCGCGCGCGCGTGGTCCAGAGGAGCGGCGCGGTGGGCCCATGGCCAGG	240
Db	181	GGCGGCGGTGGCGCGCGCGCGTGGTCCAGAGGAGCGGCGCGGTGGGCCCATGGCCAGG	240
Qy	241	ACTGTGGGCAACATCGAGGAGTGGCTGCTGAAATGTAAGAGTGACGGGCTACCCCGGGACT	300
Db	241	ACTGTGGGCAACATCGAGGAGTGGCTGCTGAAATGTAAGAGTGACGGGCTACCCCGGGACT	300
Qy	301	TTGATCCCGCTACAGATGACCTATCAAAATGAAGAGGACATCCCTCCATGTTCTCAGCT	360
Db	301	TTGATCCCGCTACAGATGACCTATCAAAATGAAGAGGACATCCCTCCATGTTCTCAGCT	360
Qy	361	ATCCGTTCTCAGCACAGCGGTGAGACATCTGCATCAACATGCTGGCTTGGCGCGGCT	420
Db	361	ATCCGTTCTCAGCACAGCGGTGAGACATCTGCATCAACATGCTGGCTTGGCGCGGCT	420
Qy	421	GACACCCCTGCTCTCAGGCAGCACCGAGTGGTGGAGGACATGTTCAATGTGAAAGCTGCTG	480
Db	421	GACACCCCTGCTCTCAGGCAGCACCGAGTGGTGGAGGACATGTTCAATGTGAAAGCTGCTG	480
Qy	481	GCCCTCAGCATCTGACACAGGAGCGCTTACAGTCCATGAAGAGCGGAAATGTGGACGAT	540
Db	481	GCCCTCAGCATCTGACACAGGAGCGCTTACAGTCCATGAAGAGCGGAAATGTGGACGAT	540
Qy	541	GGGCACATCATTAACATCAATAGCATGTCGCGCACCGAGGTACCCCTGCTGTGACC	600
Db	541	GGGCACATCATTAACATCAATAGCATGTCGCGCACCGAGGTACCCCTGCTGTGACC	600
Qy	601	CACCTTCTATAGTGCCACCAAGTATGCCGTCACTGCGCTGACAGAGGAGTCAAGCAAGAG	660
Db	601	CACCTTCTATAGTGCCACCAAGTATGCCGTCACTGCGCTGACAGAGGAGTCAAGCAAGAG	660
Qy	661	CTTGGGAGGCGCCAGACCCACATCGAGGACGCTGATCTCTCAGGTGTGGTGAGACA	720
Db	661	CTTGGGAGGCGCCAGACCCACATCGAGGACGCTGATCTCTCAGGTGTGGTGAGACA	720

Qy	721	CAATTGCGCTTCAAACTCCACGACAGGACCCCTGAGAGGAGCTGCCACCTATGAGCAA	780
Db	721	CAATTGCGCTTCAAACTCCACGACAGGACCCCTGAGAGGAGCTGCCACCTATGAGCAA	780
Qy	781	ATGAAGTGTCTCAAAACCGAGGATGTGGCGAGGCTGTATCTACGTCTCTCAGACACCCCC	840
Db	781	ATGAAGTGTCTCAAAACCGAGGATGTGGCGAGGCTGTATCTACGTCTCTCAGACACCCCC	840
Qy	841	GCAACATCCAGATTGGAGACATCCAGATGAGGCCACCGAGGAGTGAAGTGAAGTGA	900
Db	841	GCAACATCCAGATTGGAGACATCCAGATGAGGCCACCGAGGAGTGAAGTGAAGTGA	900
Qy	901	TGGAGCTCTCTCCCTCCCGACCCCTTCAATGAGTGTGCTCTCTGATTTTGG	960
Db	901	TGGAGCTCTCTCCCTCCCGACCCCTTCAATGAGTGTGCTCTCTGATTTTGG	960
Qy	961	TGTTGATTTCTGGATCAGGGATACCACTTCTCTGCACACCCGACCCAGGGGCTAGAAA	1020
Db	961	TGTTGATTTCTGGATCAGGGATACCACTTCTCTGCACACCCGACCCAGGGGCTAGAAA	1020
Qy	1021	ATTTGTTGAGATTTTATATCATCTTGTCAAAATGCTTCAAGTGTAAATGTGAAAAATG	1080
Db	1021	ATTTGTTGAGATTTTATATCATCTTGTCAAAATGCTTCAAGTGTAAATGTGAAAAATG	1080
Qy	1081	GGCTGGGGAAGGAGGTGGTGTCCCTAATGCTTTTACTTGTAACTTGTCTTGTGCCCC	1140
Db	1081	GGCTGGGGAAGGAGGTGGTGTCCCTAATGCTTTTACTTGTAACTTGTCTTGTGCCCC	1140
Qy	1141	TGGGCACTTGGCTTTTGTCTCTCAGTGTCTTCCCTTTGACATGGGAAAGGAGTTGTG	1200
Db	1141	TGGGCACTTGGCTTTTGTCTCTCAGTGTCTTCCCTTTGACATGGGAAAGGAGTTGTG	1200
Qy	1201	GCCAAAATCCCATCTTTTGTGCACTCAAGTGTGTGGCTCAGGGCTGGGGTGGCAGAGG	1260
Db	1201	GCCAAAATCCCATCTTTTGTGCACTCAAGTGTGTGGCTCAGGGCTGGGGTGGCAGAGG	1260
Qy	1261	GAGGCTTCACTTATATCTGTGTGTATCCAGGGCTCCAGACTTCTCTCTGCTGCTG	1320
Db	1261	GAGGCTTCACTTATATCTGTGTGTATCCAGGGCTCCAGACTTCTCTCTGCTGCTG	1320
Qy	1321	CCCACTGACCCCTCTCCCTTATCTATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCT	1380
Db	1321	CCCACTGACCCCTCTCCCTTATCTATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCT	1380
Qy	1381	TCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1440
Db	1381	TCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1440
Qy	1441	GGGCTGCGCCAGTGGATTTTCATGTCATTAATAAAGAAAAATCGAACCCAAAAA	1500
Db	1441	GGGCTGCGCCAGTGGATTTTCATGTCATTAATAAAGAAAAATCGAACCCAAAAA	1500
Qy	1501	AAAAA 1505	
Db	1501	AAAAA 1505	

RESULT 12

US-10-174-581-377
; Sequence 377, Application US/10174581
; Publication No. US20030017540A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3430RIC41
CURRENT APPLICATION NUMBER: US/10/174,581
CURRENT FILING DATE: 2002-06-18
PRIOR APPLICATION NUMBER: 10/052586
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063540
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063564
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063734
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063870
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066120
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066466
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066772
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069335
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069425
PRIOR FILING DATE: 1997-12-12
PRIOR APPLICATION NUMBER: 60/069870
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/068017
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086023
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/086392
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/086486
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087098
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087208
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05

```

; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088722
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088740
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088811
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088825
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088863
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089090
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653

Query Match      100.0%; Score 1505; DB 13; Length 1505;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCGGATCGGACCCAAAGCAGGTGCGCGCGCGGAGAGAGCGCGCGGTGAGTCC 60
Db 1 CGCGGATCGGACCCAAAGCAGGTGCGCGCGCGGAGAGAGCGCGCGGTGAGTCC 60

Qy 61 TCGACCCCGGTGCGGCTAGTCCAGCGAGCGCGCGCGGTGCGGCGGTGCGGCGG 120
Db 61 TCGACCCCGGTGCGGCTAGTCCAGCGAGCGCGCGCGGTGCGGCGGTGCGGCGG 120

Qy -21 CCCGCGATGAGCGGTGCGCGGACCGGTGCGGTGAGCGGCGCGGTGCGGCGGCGATC 180
Db -21 CCCGCGATGAGCGGTGCGCGGACCGGTGCGGTGAGCGGCGCGGTGCGGCGGCGATC 180

Qy 181 GCGCGCGCGGTGCGGCGCGCGCGCGGTGCGGAGGAGCTGAAAGGTGCGGCGCGCG 240
Db 181 GCGCGCGCGGTGCGGCGCGCGCGCGGTGCGGAGGAGCTGAAAGGTGCGGCGCGCG 240

Qy 241 ACTGTGGGCAACATCGAGGAGTGGCTGCTCAATGTAAAGTGCAGGCTACCCCGGACT 300
Db 241 ACTGTGGGCAACATCGAGGAGTGGCTGCTCAATGTAAAGTGCAGGCTACCCCGGACT 300

Qy 301 TTGATCCCTACAGATGTGACCTATCAAAATGAAGAGGACATCCCTCCATGTCTCAGCT 360
Db 301 TTGATCCCTACAGATGTGACCTATCAAAATGAAGAGGACATCCCTCCATGTCTCAGCT 360

Qy 361 ATCCGTTCTCAGCAAGCGGTGTAGACATCTGCATCAACAATGTCTGGCTGGCGCGGCT 420
Db 361 ATCCGTTCTCAGCAAGCGGTGTAGACATCTGCATCAACAATGTCTGGCTGGCGCGGCT 420

```

```

Qy 421 GACACCCCTGCTCTCAGGCGAGCACCAAGTGGTTGGAGGACATGTTCAATGTGAACGTGCTG 480
Db 421 GACACCCCTGCTCTCAGGCGAGCACCAAGTGGTTGGAGGACATGTTCAATGTGAACGTGCTG 480

Qy 481 GCCCTCAGCATCTGCACACGGGAAGCCCTACAGTCCATGAAGGAGCGGAATGTGACGAT 540
Db 481 GCCCTCAGCATCTGCACACGGGAAGCCCTACAGTCCATGAAGGAGCGGAATGTGACGAT 540

Qy 541 GGGCACATCAATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGTCTGTGACC 600
Db 541 GGGCACATCAATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGTCTGTGACC 600

Qy 601 CACTTCTATAGTGCCACCAAGTATGCCGTCTACCTGGCTGACAGAGGAGTCCAGCAAGAG 660
Db 601 CACTTCTATAGTGCCACCAAGTATGCCGTCTACCTGGCTGACAGAGGAGTCCAGCAAGAG 660

Qy 661 CTTGGGAGGCGCCAGACCCACATCCGAGCCACGTGCTCTCTCCAGGTGTGGTGAGACA 720
Db 661 CTTGGGAGGCGCCAGACCCACATCCGAGCCACGTGCTCTCTCCAGGTGTGGTGAGACA 720

Qy 721 CAATTGCGCTTCAAACTCCACGACCAAGGACCTGAGAAGGAGTGCCTATGAGCAA 780
Db 721 CAATTGCGCTTCAAACTCCACGACCAAGGACCTGAGAAGGAGTGCCTATGAGCAA 780

Qy 781 ATGAAGTGTCTCAAAACCGGAGGATGTGGCGAGGCTGTATCTACGTCTCAGCACCC 840
Db 781 ATGAAGTGTCTCAAAACCGGAGGATGTGGCGAGGCTGTATCTACGTCTCAGCACCC 840

Qy 841 GCACACATCCAGATTTGGAGACATCCAGATGAGGCGCCACGAGGAGTGCCTAGTACTG 900
Db 841 GCACACATCCAGATTTGGAGACATCCAGATGAGGCGCCACGAGGAGTGCCTAGTACTG 900

Qy 901 TGGAGGTCTCTCTTCCCTCCCAACCTTCAATGAGTGTGCTCTGCTCTGATTTTGG 960
Db 901 TGGAGGTCTCTCTTCCCTCCCAACCTTCAATGAGTGTGCTCTGCTCTGATTTTGG 960

Qy 961 TGTGATTTCTGGATCAGGGATACCACTTCTCTCCACACCCCGGAGGCTAGAAA 1020
Db 961 TGTGATTTCTGGATCAGGGATACCACTTCTCTCCACACCCCGGAGGCTAGAAA 1020

Qy 1021 ATTTGTTTGAGATTTTATATCATCTTGTCAAAATGCTTCAAGTTGTAATGTGAATAATG 1080
Db 1021 ATTTGTTTGAGATTTTATATCATCTTGTCAAAATGCTTCAAGTTGTAATGTGAATAATG 1080

Qy 1081 GGCTGGGGAAGAGGAGTGTGCTCCCTAATGTTTAACTTGTAACTTGTCTGTGCCCC 1140
Db 1081 GGCTGGGGAAGAGGAGTGTGCTCCCTAATGTTTAACTTGTAACTTGTCTGTGCCCC 1140

Qy 1141 TGGGCACCTTGGCCCTTGTCTGCTCTCAGTGTCTTCCCTTTCACATGGGAAAGGAGTGTG 1200
Db 1141 TGGGCACCTTGGCCCTTGTCTGCTCTCAGTGTCTTCCCTTTCACATGGGAAAGGAGTGTG 1200

Qy 1201 GCCAAAATCCCATCTTTCTTGCACCTCAACGTCTGTGGCTCAGGGCTGGGGTGGCAGAG 1260
Db 1201 GCCAAAATCCCATCTTTCTTGCACCTCAACGTCTGTGGCTCAGGGCTGGGGTGGCAGAG 1260

Qy 1261 GAGGCTTTCACCTTATATCTGTGTGTATCCAGGCTCAGACTTCTCTCTGCTGCTG 1320
Db 1261 GAGGCTTTCACCTTATATCTGTGTGTATCCAGGCTCAGACTTCTCTCTGCTGCTG 1320

Qy 1321 CCCACTGACCCCTCTCCCTTATCTATCTCTTCTCGGCTCCCGAGCCAGTCTTGCT 1380
Db 1321 CCCACTGACCCCTCTCCCTTATCTATCTCTTCTCGGCTCCCGAGCCAGTCTTGCT 1380

Qy 1381 TCTTGTCCCTCTCTGGGCTCATCCCTCCACTCTGACTCTGACTATGGCAGCAGAACCA 1440
Db 1381 TCTTGTCCCTCTCTGGGCTCATCCCTCCACTCTGACTCTGACTATGGCAGCAGAACCA 1440

Qy 1441 GGGCTGGCCCGAGTGGATTTTCATGGTGATCAATTAAGAGAGAAATCGACACCAAAAA 1500
Db 1441 GGGCTGGCCCGAGTGGATTTTCATGGTGATCAATTAAGAGAGAAATCGACACCAAAAA 1500

```



```
; Publication No. US20030017542A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C76
; CURRENT APPLICATION NUMBER: US/10/176,749
; CURRENT FILING DATE: 2002-06-20
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 377
; LENGTH: 1505
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-749-377

Query Match      100.0%; Score 1505; DB 13; Length 1505;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  CGCGGATCGGACCCCAAGCAGGTGCGCGCGCGCGCAGGAGCGCGCGGCGTCTCAGTCC 60
Db      1  CGCGGATCGGACCCCAAGCAGGTGCGCGCGCGCGCAGGAGCGCGCGGCGTCTCAGTCC 60

Qy      61  TCGACCCCGGTGTCGGGTAGTCCAGCGAGCGGCGCGCGTGGGCCCATGGCCAGG 120
Db      61  TCGACCCCGGTGTCGGGTAGTCCAGCGAGCGGCGCGCGTGGGCCCATGGCCAGG 120

Qy      121  CCGGGATGAGCGGTGCGCGACCGGTGCGCGTGGTGGGCGTGGGGGGGCATC 180
Db      121  CCGGGATGAGCGGTGCGCGACCGGTGCGCGTGGTGGGCGTGGGGGGGCATC 180

Qy      181  GCGCGCGCGGTGGCGCGCGCGTGGTCCAGCAGGAGCTGAAGGTGGTGGGCTGGCGCGCG 240
Db      181  GCGCGCGCGGTGGCGCGCGCGTGGTCCAGCAGGAGCTGAAGGTGGTGGGCTGGCGCGCG 240

Qy      241  ACTGTGGGCAACATCGAGGAGCTGGCTGCTGAATGTGAAGTGCAGGCTACCCCGGACT 300
Db      241  ACTGTGGGCAACATCGAGGAGCTGGCTGCTGAATGTGAAGTGCAGGCTACCCCGGACT 300

Qy      301  TTGATCCCTTACAGATGTGACCTATCAAAATGAAGAGGACATCCCTCTCCATGTTCTCAGCT 360
Db      301  TTGATCCCTTACAGATGTGACCTATCAAAATGAAGAGGACATCCCTCTCCATGTTCTCAGCT 360

Qy      361  ATCCGTTCTCAGCACAGCGGTAGACATCTGCATCAACAAATGCTGGCTGGCCCGGCGCT 420
Db      361  ATCCGTTCTCAGCACAGCGGTAGACATCTGCATCAACAAATGCTGGCTGGCCCGGCGCT 420

Qy      421  GACACCTGCTCTCAGCAGCAGCAGTGGTGGAGGACATGTTCAATGTGAAGTGTGCTG 480
Db      421  GACACCTGCTCTCAGCAGCAGCAGTGGTGGAGGACATGTTCAATGTGAAGTGTGCTG 480

Qy      481  GCGCTCAGCATCTGCACACGGGAGCCTACAGTCCATGAAGAGCGGAAATGTGGACAT 540
Db      481  GCGCTCAGCATCTGCACACGGGAGCCTACAGTCCATGAAGAGCGGAAATGTGGACAT 540

Qy      541  GGGCACATCATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGCTGTGACC 600
Db      541  GGGCACATCATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGCTGTGACC 600

Qy      601  CACTTCTATAGTCCACCAAGTATGCGGTGCTGCTGCTGACAGGAGCTGAGGCAAGAG 660
Db      601  CACTTCTATAGTCCACCAAGTATGCGGTGCTGCTGCTGACAGGAGCTGAGGCAAGAG 660
```

RESULT 15

US-10-176-914-377

; Sequence 377, Application US/10176914

; Publication No. US20030017543A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

```
Qy      661  CTTCCGGAGGCGCCAGAGCCACATCCGAGCCACGTCGTCATCTCTCCAGGTGTGGTGAGACA 720
Db      661  CTTCCGGAGGCGCCAGAGCCACATCCGAGCCACGTCGTCATCTCTCCAGGTGTGGTGAGACA 720

Qy      721  CAATTGCGCTTCAAACTCCACGACAAGACCTCCAGAAAGGAGCTGCCACCTATGAGCAA 780
Db      721  CAATTGCGCTTCAAACTCCACGACAAGACCTCCAGAAAGGAGCTGCCACCTATGAGCAA 780

Qy      781  ATGAAGTGTCTCAAAACCCGAGGATGTGCCCGAGGCTGTATTATCTAGCTCTCAGCACCCCC 840
Db      781  ATGAAGTGTCTCAAAACCCGAGGATGTGCCCGAGGCTGTATTATCTAGCTCTCAGCACCCCC 840

Qy      841  GCACACATCCAGATTGGAGACATCCAGATGAGGCCACCGAGGAGCTGACCTAGTACTG 900
Db      841  GCACACATCCAGATTGGAGACATCCAGATGAGGCCACCGAGGAGCTGACCTAGTACTG 900

Qy      901  TGGGAGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 960
Db      901  TGGGAGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 960

Qy      961  TGTGATTTCTGATACAGGATACCACTTCTCTCTCCACACCCCGACACAGGGGCTAGAAA 1020
Db      961  TGTGATTTCTGATACAGGATACCACTTCTCTCTCCACACCCCGACACAGGGGCTAGAAA 1020

Qy      1021  ATTTGTTTGGAGATTTTATATATCATCTTGTCAAAATGCTTCAAGTTGTAATGTGAAAATG 1080
Db      1021  ATTTGTTTGGAGATTTTATATATCATCTTGTCAAAATGCTTCAAGTTGTAATGTGAAAATG 1080

Qy      1081  GGCTGGGAAAAGAGGTGGTGTCCCTAAATGTTTAACTTGTAACTTGTCTCTGCCCC 1140
Db      1081  GGCTGGGAAAAGAGGTGGTGTCCCTAAATGTTTAACTTGTAACTTGTCTCTGCCCC 1140

Qy      1141  TGGGCACTTGGCCCTTGTCTGCTCTCAGTGTCTTCCCTTTGACATGGGAAAGGAGTTGTG 1200
Db      1141  TGGGCACTTGGCCCTTGTCTGCTCTCAGTGTCTTCCCTTTGACATGGGAAAGGAGTTGTG 1200

Qy      1201  GCCAAAATCCCACTTCTTTCGACCTCAAGTGTGTGGCTCAGGGCTGGGGTGGCAGAGG 1260
Db      1201  GCCAAAATCCCACTTCTTTCGACCTCAAGTGTGTGGCTCAGGGCTGGGGTGGCAGAGG 1260

Qy      1261  GAGGCTTCCACTTATATCTGTGTGTATCCAGGGCTCCAGACTTCTCTCTGCGCTGC 1320
Db      1261  GAGGCTTCCACTTATATCTGTGTGTATCCAGGGCTCCAGACTTCTCTCTGCGCTGC 1320

Qy      1321  CCACACTGCACTCTCTCCCTTATCTATCTCTTCTCGGTCTCCAGCCAGCTTGGCT 1380
Db      1321  CCACACTGCACTCTCTCCCTTATCTATCTCTTCTCGGTCTCCAGCCAGCTTGGCT 1380

Qy      1381  TCTTGTCCCTCTCGGGTTCATCCCTCCACTCTGACTCTGACTATGGCAGCAGAACCA 1440
Db      1381  TCTTGTCCCTCTCGGGTTCATCCCTCCACTCTGACTCTGACTATGGCAGCAGAACCA 1440

Qy      1441  GGGCTGGCCAGTGGATTTTCATGGTATCAATTAATAAAGAAAATCGCAACCAAAAAA 1500
Db      1441  GGGCTGGCCAGTGGATTTTCATGGTATCAATTAATAAAGAAAATCGCAACCAAAAAA 1500

Qy      1501  AAAAA 1505
Db      1501  AAAAA 1505
```

```

; APPLICANT: Pan,James
; APPLICANT: Smith,Victoria
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C83
; CURRENT APPLICATION NUMBER: US/10/176,914
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 377
; LENGTH: 1505
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-914-377

Query Match      100.0%; Score 1505; DB 13; Length 1505;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGATCGGACCCCAAGCAGGTGCGCGCGCGGCGGCGAGAGCGCGCGGGGTCAAGTCC 60
Db 1 CGCGATCGGACCCCAAGCAGGTGCGCGCGCGGCGGCGGCGGCGGCGGTCAAGTCC 60

QY 61 TCGACCCCGGTGTCGGGCTAGTCCAGCGGAGCGGCGGCGGCGGTGGGCCCATGSCCAGG 120
Db 61 TCGACCCCGGTGTCGGGCTAGTCCAGCGGAGCGGCGGCGGCGGTGGGCCCATGSCCAGG 120

QY 121 CCCGCGATGAGCGGTGCGCGACCGGCTGGCGCTGGTGACGCGGGGCGCTCGGGGGGCATC 180
Db 121 CCCGCGATGAGCGGTGCGCGACCGGCTGGCGCTGGTGACGCGGGGCGCTCGGGGGGCATC 180

QY 181 GCGCGGCGGTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
Db 181 GCGCGGCGGTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240

QY 241 ACTGTGGCAACATCGAGGAGTGGCTGCTGAATGTAAGAGTGCAGGCTACCCGGGACT 300
Db 241 ACTGTGGCAACATCGAGGAGTGGCTGCTGAATGTAAGAGTGCAGGCTACCCGGGACT 300

QY 301 TTGATCCCTCAGATGTGACCTATCAAAATGAAGAGGACATCTCTCCATGTTCTCAGCT 360
Db 301 TTGATCCCTCAGATGTGACCTATCAAAATGAAGAGGACATCTCTCCATGTTCTCAGCT 360

QY 361 ATCCGTTCTCAGCAGCGGTGTAGACATCTGCATCAACAATGTGGCTTGGCCCGGCT 420
Db 361 ATCCGTTCTCAGCAGCGGTGTAGACATCTGCATCAACAATGTGGCTTGGCCCGGCT 420

QY 421 GACACCTGTCTCAGGCGACGACGAGTGGTGGAGGACATGTTCAATGTGAACGTGCTG 480
Db 421 GACACCTGTCTCAGGCGACGACGAGTGGTGGAGGACATGTTCAATGTGAACGTGCTG 480

QY 481 GCCCTCAGCATCTCAGCGGAAAGCCCTACAGTCCATGAAGAGCGGAATGTGACGAT 540
Db 481 GCCCTCAGCATCTCAGCGGAAAGCCCTACAGTCCATGAAGAGCGGAATGTGACGAT 540

QY 541 GGGCAGCATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGTCTGTGACC 600
Db 541 GGGCAGCATTAACATCAATAGCATGTCTGGCCACCGAGTGTACCCCTGTCTGTGACC 600

QY 601 CACTTCTATAGTGCACCAAGTATGCGTCACTGGCTGACAGAGGAGTCTGAGGAGAG 660
Db 601 CACTTCTATAGTGCACCAAGTATGCGTCACTGGCTGACAGAGGAGTCTGAGGAGAG 660

QY 661 CTTCCGAGGCGCCAGACCCATCCGAGCCACGTGCATCTCTCCAGGTGGTGGAGACA 720
Db 661 CTTCCGAGGCGCCAGACCCATCCGAGCCACGTGCATCTCTCCAGGTGGTGGAGACA 720

QY 721 CAATTCGCTTCAAACTCCAGCAAGGACCCCTGAGAGGCGAGTGCACCTATGAGCAA 780
Db 721 CAATTCGCTTCAAACTCCAGCAAGGACCCCTGAGAGGCGAGTGCACCTATGAGCAA 780

```

```

QY 781 ATGAAGTGTCTCAAAACCGAGAGATGTGCGGAGGCTGTATATCTAGTCTCAGCACCCCC 840
Db 781 ATGAAGTGTCTCAAAACCGAGAGATGTGCGGAGGCTGTATATCTAGTCTCAGCACCCCC 840

QY 841 GCACACATCCAGATTGGAGACATCCAGATGAGGCCCGGAGCAGGTGACCTAGTGACTG 900
Db 841 GCACACATCCAGATTGGAGACATCCAGATGAGGCCCGGAGCAGGTGACCTAGTGACTG 900

QY 901 TGGGAGCTCTCTCTCCCTCCACCTTCATGGCTTGCCTCTCCTCTCTGATTTTAGG 960
Db 901 TGGGAGCTCTCTCTCCCTCCACCTTCATGGCTTGCCTCTCCTCTCTGATTTTAGG 960

QY 961 TGTGATTTCTGGATCAGGGATACCACTTCTGTCCACACCCCGAGGGGCTAGAAA 1020
Db 961 TGTGATTTCTGGATCAGGGATACCACTTCTGTCCACACCCCGAGGGGCTAGAAA 1020

QY 1021 ATTTGTTTGAGATTTTATATCATCTTGTCAAAATGCTTCAGTTGTAATGTGAAAAATG 1080
Db 1021 ATTTGTTTGAGATTTTATATCATCTTGTCAAAATGCTTCAGTTGTAATGTGAAAAATG 1080

QY 1081 GGTCTGGGAAAAGGAGGTGGTCTCCCTAAATGTTTACTTGTAACTTCTTCTGTGCCCC 1140
Db 1081 GGTCTGGGAAAAGGAGGTGGTCTCCCTAAATGTTTACTTGTAACTTCTTCTGTGCCCC 1140

QY 1141 TGGGCACTTGGCCCTTGTCTCTCAGTGTCTTCCCTTTGACATGGGAAAAGGAGTTGTG 1200
Db 1141 TGGGCACTTGGCCCTTGTCTCTCAGTGTCTTCCCTTTGACATGGGAAAAGGAGTTGTG 1200

QY 1201 GCCAAAATCCCACTCTTCTGCACTCAACGCTGTGCTCAGGCTGGGGTGGCAGAGG 1260
Db 1201 GCCAAAATCCCACTCTTCTGCACTCAACGCTGTGCTCAGGCTGGGGTGGCAGAGG 1260

QY 1261 GAGGCTTACCTTTATATCTGTGTTTATCCAGGGCTCCAGACTTCTCCTCTGCTCTGC 1320
Db 1261 GAGGCTTACCTTTATATCTGTGTTTATCCAGGGCTCCAGACTTCTCCTCTGCTCTGC 1320

QY 1321 CCCACTGACCCCTCTCCCCCTTATCTATCTCTTCTCGGCTCCCGAGCCCAAGTCTGGCT 1380
Db 1321 CCCACTGACCCCTCTCCCCCTTATCTATCTCTTCTCGGCTCCCGAGCCCAAGTCTGGCT 1380

QY 1381 TCTTGTCCCTCTCCCTGGGTCACTCCCTCCACTCTGACTCTGACTATGGCAGCAGACCA 1440
Db 1381 TCTTGTCCCTCTCCCTGGGTCACTCCCTCCACTCTGACTCTGACTATGGCAGCAGACCA 1440

QY 1441 GGGCCTGGCCCGAGTGGATTTTCATGGTGATCAATTAAGAAAAAATCCCAACCAAAAAA 1500
Db 1441 GGGCCTGGCCCGAGTGGATTTTCATGGTGATCAATTAAGAAAAAATCCCAACCAAAAAA 1500

QY 1501 AAAAA 1505
Db 1501 AAAAA 1505

```

Search completed: September 16, 2004, 12:43:53
Job time : 744 secs

Blank

EARLIER APPLICATION NUMBER: 60/055,949
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,953
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,950
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,947
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,964
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/056,360
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,684
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,984
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,954
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/058,785
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,664
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,660
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,661
EARLIER FILING DATE: 1997-09-12
NUMBER OF SEQ ID NOS: 672
SOFTWARE: Patentin ver. 2.0
SEQ ID NO 89
LENGTH: 569
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (1)
OTHER INFORMATION: n equals a,t,g, or c
US-09-227-357-89

Query March 34.1%; Score 513.6; DB 4; Length 569;
Best Local Similarity 99.4%; Pred. No. 3.3e-123;
Matches 535; Conservative 1; Mismatches 0; Indels 2; Gaps 2;
Qy 568 TTCTGGATCAGGGATACCACTTCCTGTCCACACCCGACGAGGGCTAGAAAATTGTT 1027
Db 13 TTCTGGATCAGGGATACCA-TTCTGTCCACACCCGACGAGGGCTAGAAAATTGTT 71
Qy 1028 TGAGATTTTATATCATCTTGTCAAATGCTTCAGTTGTAAATGTGAAAAATGGCTGG 1087
Db 72 TGAGATTTTATATCATCTTGTCAAATGCTTCAGTTGTAAATGTGAAAAATGGCTGG 131
Qy 1088 GAAAGGAGGTGGTCCCTAATTTTACTGTGTAACTTTCTTGTGCCCCCTGGGCAC 1147
Db 132 GAAAGGAGGTGGTCCCTAATTTTACTGTGTAACTTTCTTGTGCCCCCTGGGCAC 191
Qy 1148 TTGGCCCTTTGCTGTCTCAGTGTCTCCCTTTGACATGGGAAAGGAGTTGTGGCCAAA 1207
Db 192 TTGGCCCTTTGCTGTCTCAGTGTCTCCCTTTGACATGGGAAAGGAGTTGTGGCCAAA 251
Qy 1208 TCCCATCTTCTTGCACTCAACGTCTGTGCTCAGGCTGGGTGGGAGGAGGCT 1267
Db 252 TCCCATCTTCTTGCACTCAACGTCTGTGCTCAGGCTGGGTGGGAGGAGGCT 311
Qy 1268 TCACCTTATCTGTGTGTATCCAGGCTCCAGACTTCCTCTGCTGCCCTGCCCACTG 1327
Db 312 TCACCTTATCTGTGTGTATCCAGGCTCCAGACTTCCTCTGCTGCCCTGCCCACTG 371
Qy 1328 CACCTCTCCCTTATCTATCTCTCTCGCTCCCGACCCAGCTCTGCTGCTTCTTCTGTC 1387
Db 372 CACCTCTCCCTTATCTATCTCTCTCGCTCCCGACCCAGCTCTGCTGCTTCTTCTGTC 431
Qy 1388 CCCTCTGGGTTCATCCCTCCACTCTGACTGTGATGTGGCAGCAGAACACAGGCGCTG 1447
Db 432 CCCTCTGGGTTCATCCCTCCACTCTGACTGTGATGTGGCAGCAGAACACCA-GGCCTG 490

Qy 1448 GCCCAGTGGATTTTCATGTCATTAATAAAGAAAAATCGACCAAAAAA 1505
Db 491 GCCCAGTGGATTTTCATGTCATTAATAAAGAAAAATCGACCAAAAAA 548
RESULT 2
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-Fls
US-08-232-463-14

Query Match 4.5%; Score 67.2; DB 1; Length 7218;
Best Local Similarity 3.6%; Pred. No. 2.1e-07;
Matches 12; Conservative 209; Mismatches 117; Indels 0; Gaps 0;
Qy 1082 GCTGGGAAAGAGGTGGTGTCCCTAATTTGTTTACTTGTAACTTGTCTTGTGCCCCCT 1141
Db 1045 GCAGTCCAGGAGCTTGGATYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1104
Qy 1142 GGGCAGCTTGGCTTGTCTCTCAGTGTCTTCCCTTGTACATGGGAAGAGTTGTGG 1201
Db 1105 YY 1164
Qy 1202 CCAAAATCCCATCTTCTTGACCTCAACGCTCTGTGGCTCAGGCTGGGGTGGCAGGG 1261
Db 1165 YY 1224
Qy 1262 AGGCCTTCACCTTATCTGTGTGTATCCAGGCTCCAGACTTCCTCTCTGCTGCTGCC 1321
Db 1321 YY 1380

Db 8129 TGCTGGTCAACAACGCGGACGAAACGCGGTGCGGTCACTCCACCTCACGACGAGC 8070
Qy 449 GTTGAAGGACATTTCAATGTGAACTGCTGGCCCTCAGCATCTGCACACGGAAAGCCT 508
Db 8069 TGTGGCAGGACGTATCGACACCAACCTCACAGCGTCTTCCGCGTCAACCGCGAGGTCC 8010
Qy 509 ACCAGTCCATGAAGGACGGAATGTGACGATGGGCACATCATTAACATCAATAGCATGT 568
Db 8009 TCACACCGCGGCGATGAGAGCGCGCGGACGCGCGGATCATCAGGTGCGCTCCACCG 7950
Qy 569 CTGGCCACCGAGTGTACCCCTGTCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCG 628
Db 7949 GCGGCAAGCAGGCGTCCGCTGGCGCGCCCTACT-----CGGCTTCCAAAGCGCGG 7896
Qy 629 TCACTGGCTGACAGAGGACTGAGGCAAGAGCTTCGGAGGCGCCAGACCCACATCCGAG 688
Db 7895 TCATCGGCTTCAACAAGCGCTGGCCAAAGAACT-----CGCCACACCGCGCACCGG 7842
Qy 689 CCAGTGCATCTTCCAGGTGTGTGGAGACACAAATTCGCTTCAAACTCCACGACAAGG 748
Db 7841 TCAACGCGCTTCCGCGGCTACGTGAGAGCGCGATGGCGGTGCGGGTCCGCCAGGCGT 7782
Qy 749 ACCGTGAGAGGACGCTGCCACCTATGAGCAAAATG 783
Db 7781 ACGCGCGACCTGGCGGACCAACCGAGGAGGACGTG 7747

RESULT 5

US-08-474-933-1/c
; Sequence 1, Application US/08474933
; Patent No. 5866410
; GENERAL INFORMATION:
; APPLICANT: Ryan, Michael J.
; APPLICANT: Lotvin, Jason A.
; APPLICANT: Strathy, Nancy
; APPLICANT: Fantini, Susan E.
; TITLE OF INVENTION: Cloning of the biosynthetic pathway for
; TITLE OF INVENTION: chlortetracycline and tetracycline formation and cosmid
; TITLE OF INVENTION: useful therein
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: American Cyanamid Company
; STREET: One Cyanamid Plaza
; CITY: Wayne
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07470
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,933
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/125,468
; FILING DATE: 22-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Tsevdos, Estelle J
; REGISTRATION NUMBER: 31,145
; REFERENCE/DOCKET NUMBER: 31,255-02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201)831-3241
; TELEFAX: (201)831-3305
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30001 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)

US-08-474-933-1

Query Match 4.0%; Score 60.2; DB 2; Length 30001;
Best Local Similarity 47.1%; Pred. No. 2.5e-05;
Matches 299; Conservative 0; Mismatches 318; Indels 18; Gaps 3;
Qy 149 TGGCGCTGTGTGAGCGGGGGCCCTCGGGGGGCATCGGGCGCGCGCTGGCCCGGCGCCCTGGTCC 208
Db 8363 TCGCCCTGTGTGACCGGGGCGACCGAGCGGCATCGGACTGGCCAGCGCGCGGTGGCCC 8304
Qy 209 AGCAGGAGCTGAAGGTGGTGGGTGGCCCGCGACTGTGGGCAACATCAGAGAGCTGGCTG 268
Db 8303 GCGCGGCGCTGAGCGTCTTTCATCTGCGCCCGCACCGCGCAGGCGTGCCTCCGAGACGCTG 8244
Qy 269 CTGNACTGAAGATGTCAGGCTACCCCGGAGCTTTCAGCTATCCCTTCTCAGCATCGAGTGTGACCTATCAA 328
Db 8243 CCGCGCTGCGCGCGGAGGGGCTTGGAGGTGAGCGGCTTCGGCGCGAGCTCGCTCCCGGC 8184
Qy 329 ATGAAGAGGACATCTCTCCATGTTCTCAGCTATCCCTTCTCAGCATCGAGCGGTGTAGACA 388
Db 8183 CGGACGTGAGCAGGCTCGTCCGGCGCGGGTGGACCGCTT-----CGGCGCGATCGCG 8130
Qy 389 TCTGCATCAACAATGTGGGTGGCCCGCGCTGACACCTCTCTCAGGCGAGCAGCTG 448
Db 8129 TGTGTGTCACACACGCGCGAGCAACGCGCGTGGCGTCACTCCACCTCACCACGAGC 8070
Qy 449 GTTGAAGGACATGTTCAATGTGAACGTGTGGGCCCTCAGCATCTGCACACGGAAGCCT 508
Db 8069 TGTGGCAGGAGTGTATCGACCACTCACCAGCGTTCGCGGTACCGCGAGGTCC 8010
Qy 509 ACCAGTCCATGAAGGAGCGGAATGTGACGATGGACGATGGGCACATCATTAACATCAATAGCATGT 568
Db 8009 TCACCACCGCGCATGAGCGCGCGCGGCGCGGCGGATCATCAGGTTCGCTCCACCG 7950
Qy 569 CTGCCACCGAGTGTATACCTCTGTGTGACCGCTCTATAGTGCACCGAGTATGCCG 628
Db 7949 GCGGCAAGCAGGCGTGTCCCGTGGCGCGCCCTACT-----CGGCTTCCAAAGCGCGG 7896
Qy 629 TCACTGCGCTGACAGAGGAGCTGAGGCAAGAGCTTCGGAGGCGCCAGACCCACATCCGAG 688
Db 7895 TCATCGGCTTCAACAAGCGCTGGCCAAAGAACT-----CGCCACACCGCGCACCGG 7842
Qy 689 CCAGTGCATCTCTCAGGTGTGTGGAGACACAAATTCGCTTCAAACTCCACGACAAGG 748
Db 7841 TCAACGCGCTCTGCCCGGCTACGTGAGAGCGCGATGGCGGTGCGGGTCCGCCAGGCGT 7782
Qy 749 ACCGTGAGAGGAGCTGCCACCTATGAGCAAAATG 783
Db 7781 ACGCGCGACCTGGCGGCGACCAACCGAGGAGGACGTG 7747

RESULT 6

US-09-252-991A-15912/c
; Sequence 15912, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 15912
; LENGTH: 732
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15912

Query March	3.9%;	Score 58.4;	DB 4;	Length 732;
Best local Similarity	46.2%;	Pred. No.1.4e-05;		
Matches 234;	Conservative	0;	Mismatches 266;	Indels 6;
Gaps	1;			

Qy	23	CGGCGGCGGCGGAGAGCGGCGGCGT	CAGCTCTCGACCCCGGTGTCGGGCTAGT	82
Db	524	CAGCCACGCGACGCGGCGGACTGT	CGGCTCTCCCAAGCCCTGCGCATACACCT	465
Qy	83	CCAGCGAGCGGACGGCGGCGTGGGCCCAT	TGGCAGCGCCGGCATGGAGCGTGGCGCG	142
Db	464	ACAAGAACAATCAGAGGTGAAGCATG	CACGCTACGTCGTCGACGATTTCTCCCTGCTT	405
Qy	143	ACCGCTGGCGCTGGTGACGGGGCCCTC	GGGGGCGATCGGGCGCGCGTGGCCCGGGGCC	202
Db	404	CCCGGGTCGCGCTGGTGACCGCCCGGAC	CGCGCATCGGCCGGGCGATCGCCCTCGCCC	345
Qy	203	TGGTCCAGCAGGGACTGAAGTGTGGGCT	GGCGCCGCACTGTGGGCAACATCGAGAGC	262
Db	344	TGGCGGGCGCGCGCGACGTGGCGTGG	CGGACCTCGATCCGACGGTCGCGAGGAA	285
Qy	263	TGGCTGCTGAATGTAAAGAGTSCAGGCT	ACCCCGGAGCTTTGATCCCCCTACAGATGTGACC	322
Db	284	CCGCGCGCGGATCCGAGCGCTCGT	CGCGCAGCC-----TGGCCCTTGGGGTCAAG	231
Qy	323	TATCAATTAAGAGGACATCTCTCCAT	TGTTCTCAGCTATCGGTTCTCAGCAGACGGGT	382
Db	230	TCAGCAGCGGCGACAGCGGTCCGGGCG	ATGGTTCGAGCGGGTCCGCACGGAATTCGGCCGGC	171
Qy	383	TAGACATCTGCATCAAAATGCTGGCT	TGGCCCGGCTGACACCCCTGCTCTCAGGAGCA	442
Db	170	TCGACGTGGCGGTGAACAAATGCCGG	GTGATCAGCATCCGCAAGGTGCGCGAACTCAGCC	111
Qy	443	CCAGTGGTTGGAGGACATGTTCAAT	GTGAACGTGCTGGCCCTCAGCATCTGCACACGGG	502
Db	110	TCGCCGACTGGACCGGCTGATGAAC	TGATCAACGCCCGCGGGGTCTCTGTGTCCAGG	51
Qy	503	AAGCTACCACTCAATGAAGAGCGG	528	
Db	50	CCGAACGTGGCGCTGATCAGCGCAG	25	

```

RESULT 7
US-09-252-991A-15779
; Sequence 15779, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 15779
; LENGTH: 2073
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15779

```

	Query Match	3.9%	Score 58.4	DB 4	Length 2073
	Best Local Similarity	46.2%	Pred. No. 2.2e-05		
	Matches 234	Conservative 0	Mismatches 266	Indels 6	Gaps 1
QY	23	CGCGCGCGCGGACGAGCGGCGCGGTCTCTCGACCCCGGTGCGGCGTAGT	82		
Db	153	CAGCCACGGCGACGGAGGGCGGACTGTCGGTCTCTCCAAAGCCCTCCGCATACACCT	212		
QY	83	CCAGCGAGGCGGACGCGGCGCGTGGGCGCCATGGCCAGGCGCGGTGGCGCG	142		

Db	213	ACAAGAACAATACAGAGGTGAAGCATGCAACGCTACGTCGTCCGACGATTTCTCCCTGTCTT	272
Qy	143	ACCGGCTGGCGCTGCTGACCGGGGCGCTCGGGGGGCATCGGGCGCGCGTGGCCCGGGCCC	202
Db	273	CCCGGCTGCGCTGCTGACCGGGCCCGGACGCGGCATCGGCCGGGGCATCGCCCTCGCCC	332
Qy	203	TGCTCCACAGGAGCTAAGGTGCTGGGCTGCGCCCGCACCTGTGGGCAAATCGAGGAGC	262
Db	333	TGGCGCGCGCGCGCGACGCTGGCGGTGGCGACTCGATCCGCAGGTGCCGAGGAAA	392
Qy	263	TGGCTGCTGAATGTAAGAGTGCAGGCTACCCCGGACTTTGATCCCCCTACAGATGTACC	322
Db	393	CCGCGCGCGCATCCGACGCTCGGTGCGCGCAG-----CCTGGCCCTTGGGGTCCGACG	446
Qy	323	TATCAATGAAGGACATCTCTCCATGTTCTCAGCTATCCGTTCTCAGGACAGCGGTG	382
Db	447	TCAGCGACGGCGACAGCGTTCGGGCGATGGTTCGAGCGGGTTCGCACGGAGTTCGGCGCGC	506
Qy	383	TAGACATCTGCATCAACAATGCTGGCTTGGCCGGCGCTGACACCCCTGCTCTCAGGAGCA	442
Db	507	TCGACGTGGCGGTGTAACAATGCCGGGTGATCAGATCCGCAAGGTGCGCGAACTCAGCC	566
Qy	443	CCAGTGGTTGAAGGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGG	502
Db	567	TCGCGGACTGGACCGGGTGATGAACGTCAACGCCCGCGCGGTCTCTGTGTCGACGG	626
Qy	503	AAGCTTACCAGTCCATGAAGAGCGG	528
Db	627	CCGAACCTGCCGCTGATCAGGCGCAG	652

RESULT 8
 US-09-252-991A-13796
 ; Sequence 13796, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 13796
 ; LENGTH: 1419
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-13796

	Query Match	3.9%;	Score 58.2;	DB 4;	Length 1419;
	Best Local Similarity	49.4%;	Pred. No. 2.1e-05;		
	Matches 212;	Conservative 0;	Mismatches 208;	Indels 9;	Gaps 2;
Qy	101	GGCGTGGGCGCATGCGCCAGGCGCGGCATGAGCGGTGGCGCGACCGCGCTGGCGCTGCTGA	160		
Db	702	GGGGATAGCCCATGTCTCAAGACCACCTGTTGCACCTCGAGCGGAAGATGCGCTTTCGTTT	761		
Qy	161	CGGGGGGCTCGGGGGGATCGGCGCGCCCTGGCGCGGCCCTTGTCACAGGACTGA	220		
Db	762	CCGGCGTCAGCCGCGGATCGGGAGGCGATAGCCAAGCTGCTGGCCAGCAGGCGGCC	821		
Qy	221	AGGTGTGGGCTCGCCCGCACTGTGGGCAACATCGAGGAGCTGGCTGCTGAATGTAAGA	280		
Db	822	ACGTGATGCTGTGAGCCCGCAAGATCGATGGCTGCCAGGCGTTCGCGACGCGATCACCG	881		
Qy	281	GTGAGGCTACCCGGGACTTTGATCCCTTACAGATGTGACCTTACAAATGAAGAGACA	340		
Db	882	CGAGGGCGGCAAGGCCACCGCATCGCTGCC-----ACATCGGCGAGATGGAGCAGA	935		

COUNTRY: U.S.A.
 ZIP: 20004
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/726,614
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/790,462
 FILING DATE: 29-JAN-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: D. Douglas Price
 REGISTRATION NUMBER: 24,514
 REFERENCE/DOCKET NUMBER: 1418/P57452US2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202 638-6666
 TELEFAX: (202) 39305350
 TELEX: RCA 248593 IDEA UR
 INFORMATION FOR SEQ ID NO: 22:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 744 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 US-09-726-614-22

Query Match 3.8%; Score 56.8; DB 4; Length 744;
 Best Local similarity 44.9%; Pred. No. 3.7e-05;

RESULT 12
US-09-385-040-22/c
; Sequence 22, Application US/09385040
; Patent No. 6589775
; GENERAL INFORMATION:
; APPLICANT: Jensen, Susan E
; APPLICANT: Aidoo, Kwamena A
; APPLICANT: Paradkar, Ashish S
; TITLE OF INVENTION: DNA SEQUENCE ENCODING ENZYMES OF CLAVULANIC ACID
; TITLE OF INVENTION: BIOSYNTHESIS
; FILE REFERENCE: 09/385,040
; CURRENT APPLICATION NUMBER: US/09/385,040
; CURRENT FILING DATE: 1999-08-30
; PRIOR APPLICATION NUMBER: US 08/790,462
; PRIOR FILING DATE: 1997-01-29
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 744
; TYPE: DNA
; ORGANISM: Streptomyces clavuligerus
US-09-385-040-22

	Query Match	3.8%	Score 56.8;	DB 4;	Length 744;
	Best Local Similarity	44.9%;	Pred. No.	3.7e-05;	
	Matches	263; Conservative	0;	Mismatches 317;	Indels 6; Gaps 1;
QY	149	TGCGCTGTGTACGGGGGGCGTCCGGGGGGCATCGCGCGCGCGGCCCGGCCTTGTC	208		
Dd	719	TGCGCTGTATCAGGGCGGAGATTGGGGCATCGCGAGGCACGGCCCGGCTTGGCGG	660		

RESULT 13
US-09-385-028-13/c
; Sequence 13, Application US/09385028
; Patent No. 6232106
; GENERAL INFORMATION:
; APPLICANT: Susan E. Jensen
; APPLICANT: Kwamena A Aidoo
; APPLICANT: Ashish S. Paradkar
; TITLE OF INVENTION: DNA Sequence Encoding Enzymes of Ciavulanic
; Patent No. 6232106
; TITLE OF INVENTION: Acid Biosynthesis
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN,PPLC
; STREET: The Jenifer Bulding, 400 Seventh Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patencin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/385,028
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/790,462
; FILING DATE: 29-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: D. Douglas Price
; REGISTRATION NUMBER: 24,514
; REFERENCE/DOCKET NUMBER: 1418/P57452US

```

/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: U.S.A.
/ ZIP: 20004
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30 (BPO)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/726,614
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/790,462
/ FILING DATE: 29-JAN-1997
/ ATTORNEY/AGENT INFORMATION:
/ NAME: D. Douglas Price
/ REGISTRATION NUMBER: 24,514
/ REFERENCE/DOCKET NUMBER: 1418/P57452US2
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202 638-6666
/ TELEFAX: (202) 39305350
/ TELEX: RCA 248593 IDEA UR
/ INFORMATION FOR SEQ ID NO: 13:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 11604 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ US-09-726-614-13

Query Match      3.8%; Score 56.8; DB 4; Length 11604;
Best Local Similarity 44.9%; Pred.No.0.00012;
Matches 263; Conservative 0; Mismatches 317; Indels 6; Gaps 1;

QY    149 TGGCGCTGTTGACGGGGGCCTCGGGGGGCATCGGCGGGCGTGGCCCGGGCCTGTGGTCC 208
DB    11308 TCGCGTCTACAGGCGCGAGCTCGGGCATCGGAGGCCACGGCCCGCGCCTGGCGG 11249

QY    209 AGCAGGACTGAAGTGTTGGGTGCGCCCGCACTGTGGGCACAATCGAGGAGCTGGGTG 268
DB    11248 CCAGGCGCGCGCGTGGCCATCGCGCGCGCGGGTTCAGAAGCTGCCGCCCTGGGTG 11189

QY    269 CTGAATTGAAGTAGTCAGGCTACC CGGACTTTGATNCCCCTACAGATGTGACCTATCAA 328
DB    11188 ACAGCTGACCGCGCGCGG-----GCAAGGTCCATGCTCTCGAACTCGACGTCGCG 11135

QY    329 ATGAAGAGGACATCTCTCCATGTTCTCAGCTATCGTCTCAGCACAGCGGTGTAGACA 388
DB    11134 ACCGCGAGGGGTGNACGCGCGGTCCCTCCACGTGAGGCGCTGGCGCGCCTCGACA 11075

QY    389 TTGTGATCAACAATGCTGGCTTTGGCCCGCCTGACACCCCTGCTCTCAGGCACGACCACTG 448
DB    11074 TCCTCGTCAACAACGCGGGATCATGCTGCTCGGCCGTGGAGGACGCCGACACCACCG 11015

QY    449 GTTGAAGGACATGTTCAATGTGAAGTGTGGCCCTGAGCATGTGCACAGGGAAGCCT 508
DB    11014 ACTGGACCGGATGATCGACACCAATCTCTCGGGCTGTATGATGACCGGCGGCC 10955

QY    509 ACCAGTCCAATGAAGGAGCGGAATGTGGACATGTGGCACATCATTAACATCAATAGCATGT 568
DB    10954 TTCCCCATCTGTCGCACAGAGGCAACGTGTGCAGATGCTTCGATCGCGGCGCGG 10895

QY    569 CTGGCCACCGAGTGTTA CCCCTGTCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCG 628
DB    10894 TGAACGTTCGCAACGCGGCGCTCTACCAGGCCACGAAGTTCGGTGTGAACCGCTTCAGCG 10835

QY    629 TCAC TGCGCTGACAGAGGACTGAGGCAAGAGCTTCGGAGGCCACAGACCACATCCGAG 688
DB    10834 AGACGCTGCGCCAGGAGGTACCGAGCGCGGGTTCGGGTGCTGCTATCGAGCGCGGCA 10775

```

Search completed: September 16, 2004, 12:21:20
Job time : 150 secs

A handwritten mark, possibly a signature or initials, is written over a horizontal line. The mark consists of a large, stylized 'J' or 'G' shape with a small crossbar.